

**GENÓMICA COMPARATIVA DE LOS LOCI DE HORMONA
DEL CRECIMIENTO DE HUMANO Y CHIMPANCÉ.**

Aprobación de la tesis:



Dr. HUGO ALBERTO BARRERA SALDAÑA

Director de Tesis



Dr. GERARDO RAYMUNDO PADILLA RIVAS

Co-Director



Dr. JORGE ÁNGEL ISIDRO ASCACIO MARTÍNEZ

Co-Director

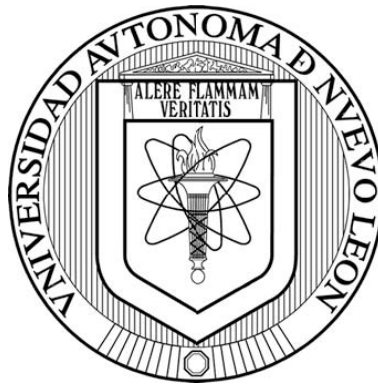


Dr. DIONICIO A. GALARZA DELGADO

Subdirector de Estudios de Posgrado

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE MEDICINA



GENÓMICA COMPARATIVA DE LOS LOCI DE HORMONA
DEL CRECIMIENTO DE HUMANO Y CHIMPANCÉ.

Presentado por

BIOQUÍMICO ANTONIO ALÍ PÉREZ MAYA

Como requisito parcial para obtener el Grado de
MAESTRIA EN CIENCIAS CON ESPECIALIDAD EN
BIOLOGÍA MOLECULAR E INGENIERÍA GENÉTICA

AGOSTO, 2008

**GENÓMICA COMPARATIVA DE LOS LOCI DE HORMONA
DEL CRECIMIENTO DE HUMANO Y CHIMPANCÉ.**

Aprobación de la tesis:

Dr. HUGO ALBERTO BARRERA SALDAÑA
Director de Tesis

Dr. GERARDO RAYMUNDO PADILLA RIVAS
Co-Director

Dr. JORGE ANGEL ASCACIO MARTÍNEZ
Co-Director

Dr. DIONICIO A. GALARZA DELGADO
Subdirector de Estudios de Posgrado

ÍNDICE GENERAL	PÁGINA
LISTA DE CUADROS	i
LISTA DE FIGURAS	ii
NOMENCLATURA	v
CAPÍTULO I. INTRODUCCIÓN	
1.1. Introducción a la secuenciación y comparación de genomas	1
1.2. Antecedentes	2
1.2.1. Estado actual de los Proyectos de secuenciación del Genoma Humano y del Chimpancé	2
1.2.2. Genómica comparativa como herramienta de estudio	5
1.2.3. Comparaciones genómicas humano-chimpancé	7
1.2.4. Genómica comparativa de <i>loci</i> definidos: el proyecto ENCODE	8
1.2.5. Evolución del <i>locus</i> GH en primates	10
1.2.6. Anatomía y fisiología del <i>locus</i> HGH	12
1.2.6.1. Estructura del <i>locus</i> HGH	12
1.2.6.2. Papel funcional de las hormonas del <i>locus</i> HGH	15
1.2.7. Regulación de la expresión de los genes del <i>locus</i> GH en el Humano	16
1.2.8. Estado del <i>locus</i> GH del chimpancé al comenzar este trabajo	20
1.3. Justificación del trabajo	21
CAPÍTULO II. OBJETIVOS	
2.1. Objetivo general	22
2.2. Objetivos específicos	22
CAPÍTULO III. ESTRATEGIA GENERAL	23
CAPÍTULO IV. MATERIAL Y MÉTODOS	
4.1. Material	24

3.1.1. Origen de los reactivos	24
3.1.2. Material biológico	25
4.2. Equipo	25
3.2.1. Programas computacionales	26
4.3. Métodos	27
4.3.1. Extracción de ADN del BAC	27
4.3.2. Amplificación de genes y regiones intergénicas	27
4.3.2.1. Oligonucleótidos diseñados en este trabajo	27
4.3.2.2. Condiciones de amplificación de los genes del chimpancé	29
4.3.2.3. Condiciones de amplificación de las regiones intergénicas del chimpancé	30
4.3.3. Clonación de los genes y regiones intergénicas	31
4.3.3.1. Preparación de células calcio competentes	31
4.3.3.2. Transformación de células calcio competentes	32
4.3.3.3. Tamizaje de las clonas recombinantes portadoras de los genes GH/CSH	32
4.3.3.4. Transformación de células electrocompetentes	33
4.3.3.5. Extracción del DNA plasmídico	33
4.3.3.6. Selección y análisis de las clonas recombinantes portadoras de los fragmentos del <i>locus</i> GH del chimpancé amplificados por PCR larga	34
4.3.4. Análisis de las secuencias	34
4.3.4.1. Secuencias utilizadas en este estudio	34
4.3.4.2. Identificación de las secuencias generadas	35
4.3.4.3. Ensamblaje de los <i>loci</i> GH	35
4.3.4.4. Análisis de las secuencias nucleotídicas de los genes	35
4.3.4.5. Análisis de las proteínas	36
4.3.4.6. Análisis del promotor proximal	36
4.3.4.7. Análisis de los elementos reguladores del control transcripcional	36
4.3.4.8. Identificación de los elementos repetitivos	36

CAPÍTULO V. RESULTADOS

5.1. Iniciando el análisis de los datos sobre el <i>locus</i> GH del genoma del chimpancé	37
5.2. Actualización del ensamblaje del <i>locus</i> GH propuesto	38
5.3. Aclarando dudas sobre los genes	40
5.4. Las regiones intergénicas como claves en el arreglo de los genes	43
5.5. Un nuevo enfoque para verificar el ensamblaje del <i>locus</i>	47
5.6. Confirmación del arreglo de cinco genes revela la organización del <i>locus</i> GH en el chimpancé	52
5.7. Ensamblaje final del <i>locus</i> GH	55
5.8. Análisis de las proteínas	55
5.9. Comparaciones de los genes placentarios de los <i>loci versus</i> los hipofisarios	57
5.10. Comparaciones de las proteínas maduras GH/CSH de humano y chimpancé	59
5.11. Análisis filogenético de los genes GH/CSH de humano y chimpancé	61
5.12. Comparaciones de las regiones intergénicas de humano y chimpancé	62
5.13. Un vistazo a reguladores de la transcripción	63
5.13.1. El promotor proximal	63
5.13.2. El inhibidor hipofisario o elemento P	64
5.13.3. El potenciador placentario o enhancer	65
5.13.4. Elementos de respuesta a glucocorticoides	66

CAPÍTULO VI. DISCUSIÓN

6.1. Obtención de los genes GH/CSH y sus regiones intergénicas	68
6.2. Organización de los genes GH/CSH en el chimpancé	68
6.3. Comparaciones de los genes GH/CSH de humano y chimpancé	70
6.4. Análisis de las proteínas	71
6.5. Análisis de los elementos reguladores	72
6.6. Análisis del promotor proximal	73
6.7. El elemento P	74
6.8. El potenciador	74

6.9. Elemento de respuesta a glucocorticoides	75
6.10. Regiones intergénicas	76
6.11. Elementos repetitivos	76
6.12. El valor de la genómica comparativa	77
6.13. Perspectivas	77
 CAPÍTULO VII. CONCLUSIONES	 79
 CAPÍTULO VIII. BIBLIOGRAFÍA	 80
 CAPÍTULO IX. ANEXOS	 88

LISTA DE CUADROS	PÁGINA
Cuadro 1. Iniciadores utilizados para amplificar por PCR los genes del <i>locus</i> GH del chimpancé.	27
Cuadro 2. Iniciadores utilizados para generar por PCR fragmentos con traslapes que permiten abarcar todo el <i>locus</i> GH del chimpancé.	28
Cuadro 3. Condiciones de amplificación de los genes GH/CSH del chimpancé.	30
Cuadro 4. Composición de las mezclas de reacción de la PCR larga.	31
Cuadro 5. Programa de amplificación de la PCR larga.	31
Cuadro 6. Secuencias utilizadas en este estudio.	34
Cuadro 7. Comparación entre genes GH/CSH de nuestro laboratorio y del CSGCh.	38
Cuadro 8. Cambios aminoacídicos entre las proteínas inferidas en este trabajo y sus contrapartes en el humano.	57

LISTA DE FIGURAS	PÁGINA
Figura 1. Evolución de los primates.	11
Figura 2. Anatomía del <i>locus</i> hGH.	13
Figura 3. Distribución de elementos Alu en el <i>locus</i> hGH.	14
Figura 4. Región de control del <i>locus</i> hGH.	17
Figura 5. Localización de sitios que participan en la regulación de la expresión de los genes del <i>locus</i> hGH.	18
Figura 6. Detalles de los sitios de regulación de la expresión de los genes del <i>locus</i> hGH.	19
Figura 7. Estrategia general.	23
Figura 8. Iniciadores utilizados para amplificar por PCR los genes del <i>locus</i> GH del chimpancé.	28
Figura 9. Iniciadores utilizados para amplificar el <i>locus</i> GH del chimpancé.	29
Figura 10. Ensamblaje preliminar de la región correspondiente al <i>locus</i> GH del chimpancé.	37
Figura 11. Actualización de las secuencias en la región génica del <i>locus</i> GH del chimpancé.	39
Figura 12. Verificación de la integridad del ADN del BAC y de la amplificación de los genes del <i>locus</i> GH del chimpancé.	41
Figura 13: Tamizaje mediante PCR en colonia de vectores recombinantes portadores de GHs.	41
Figura 14. Caracterización de los productos amplificados por digestión con <i>Pst</i> I.	42
Figura 15. Caracterización de los productos amplificados por digestión con <i>Bam</i> H I.	43
Figura 16. Amplificación de las regiones intergénicas del <i>locus</i> GH del chimpancé.	44
Figura 17. Arreglos alternativos de genes propuestos para el <i>locus</i> GH del chimpancé.	45
Figura 18: Tamizaje de las clonas recombinantes obtenidas para el FGTO10K.	46

Figura 19. Caracterización del producto amplificado obtenido para FGTO10K.	47
Figura 20. Caracterización con <i>Avr</i> II de las clonas recombinantes obtenidas para el FGTO10K.	48
Figura 21. Caracterización con <i>Bst</i> E II de las clonas recombinantes obtenidas para el FGTO10K.	48
Figura 22. Caracterización con <i>Pst</i> I de las clonas recombinantes obtenidas para el FGTO10K.	49
Figura 23. Representación de las regiones amplificadas con el par de iniciadores 06-08.	49
Figura 24. Caracterización de los productos amplificados por digestión con <i>Stu</i> I.	50
Figura 25: Tamizaje de las clonas recombinantes obtenidas para el FGTO8K.	51
Figura 26. Caracterización de las clonas obtenidas para FGTO8K.	51
Figura 27. Representación de las regiones amplificadas con el par de iniciadores 01-10 en los dos arreglos de genes propuestos para el <i>locus</i> GH del chimpancé.	52
Figura 28. Caracterización del producto amplificado de las intergénicas largas con flancos en genes.	53
Figura 29. Caracterización de los amplicones obtenidos para otras regiones del <i>locus</i> .	54
Figura 30. Organización del <i>locus</i> GH.	55
Figura 32. Alineamiento de las secuencias aminoacídicas.	56
Figura 33. Análisis comparativo de los genes GH/CSH parálogos, con respecto a las GHN.	58
Figura 34. Análisis comparativo entre genes ortólogos.	59
Figura 35. Análisis comparativo de las proteínas maduras de humano y chimpancé respecto a GHN.	60
Figura 36. Análisis comparativo entre proteínas ortólogas.	61
Figura 37. Árbol filogenético de los genes GH/CSH de humano y chimpancé.	62
Figura 38. Análisis comparativo entre regiones intergénicas que se presumen ortólogas.	62

Figura 39. Alineamiento del promotor proximal.	63
Figura 40. Alineamiento del inhibidor hipofisiario.	64
Figura 41. Alineamiento del potenciador placentario.	65
Figura 42. Identificación del GRE en el alineamiento de los genes GH/CSH de humano y chimpancé.	66

NOMENCLATURA

µg	Microgramo
µM	Concentración micromolar
°C	Grados centígrados
ADN	Acido desoxirribonucleico
dATP	Desoxiadenosina trifosfato
dNTPs	Desoxirribonucleósidos trifosfato
h	Horas
min	Minutos
mM	Concentración milimolar
ml	Mililitro
ng	Nanogramos
µl	Microlitros
pb	Pares de bases
PCR	Reacción en Cadena de la Polimerasa
pH	Potencial de hidrógeno
Taq	DNA polimerasa de <i>Thermus aquaticus</i>
X	Veces la concentración
UV	Luz Ultravioleta
EDTA	Acido etilendiaminotetra-acético
r.p.m.	Revoluciones por minuto
V	Voltios
GH	Hormona del crecimiento
U	Unidades de enzima
cbp	Cuanto basta para
DO	Densidad óptica
PA	Producto amplificado
nm	Nanómetro

AMPc	Adenosina monofosfato cíclico
BAC	Cromosoma artificial bacteriano
CSH	Hormona somatomamotropina coriónica
GH	Hormona del crecimiento
IGF-1	Factor de crecimiento semejante a la insulina tipo I
kpb	Kilo par de base
kV	Kilo volts
LCR	Región de Control del <i>locus</i>
M	Concentración molar
mg	Miligramos
MNM	Monos del nuevo mundo
MVM	Monos del viejo mundo
MYA	Millones de años atrás
ng	Nanogramos
RNA	Ácido ribonucleíco
SNP	Polimorfismo de un solo nucleótido
TE	Solución amortiguadora de Tris y EDTA
X	Número de veces la concentración
μF	Micro Faradios
Ω	Ohms

CAPÍTULO I: INTRODUCCIÓN

1.1. Secuenciación y comparación de genomas.

Desde principios de la década de 1990, muchos laboratorios se prepararon para iniciar el desciframiento de los genomas de especies que han constituido los principales modelos de investigación. El avance en las técnicas de secuenciación y el desarrollo de soportes bioinformáticos hicieron posible que para el año 1995 el primer genoma microbiano, el de *Haemophilus influenzae* (1,800,000 pares de bases o 1.8 Mb), fuese completamente secuenciado por el equipo de Craig Venter en TIGR (The Institute for Genomic Research) (Fleischmann *et al.*, 1995).

Desde entonces numerosos proyectos de secuenciación han sido completados. Los hitos más relevantes de esta nueva era han sido la secuenciación del genoma de la levadura como el primer eucariota inferior (*Saccharomyces cerevisiae*, 1997), de la mosca de la fruta como el primer invertebrado (*Drosophila melanogaster*, 2000), de la primera planta (*Arabidopsis thaliana*, 2000) y por supuesto de nuestra especie (*Homo sapiens*, 2003). Hasta el momento han sido secuenciados los genomas de más de 500 organismos y alrededor de 3000 nuevos proyectos de secuenciación se encuentran en curso (GOLD. News, 2008).

Con el Proyecto del Genoma Humano (PGH) se buscó elucidar el orden de los aproximadamente tres mil millones de nucleótidos que conforman el genoma humano, localizar en estos los genes e iniciar el análisis del resto del material hereditario de nuestra especie (Lander *et al.*, 2001). Este proyecto constituye el primer gran esfuerzo coordinado internacionalmente en la historia de la Biología. Comenzó formalmente en 1990 y aunque originalmente fue planeado para realizarse en 15 años, los rápidos avances tecnológicos acortaron en dos años su fecha de culminación (Roberts *et al.*, 2001).

Proyectos Genómicos es el término genérico con el que se designa al nuevo enfoque de la Biología en el cual, a partir de la estructura primaria del material hereditario o genoma de una serie de organismos modelo, es posible acelerar la generación del conocimiento sobre los procesos biológicos, fisiológicos y patológicos de los seres vivos, así como traducir

dicho nuevo conocimiento en aplicaciones técnicas y comerciales en ámbitos tales como el diagnóstico y tratamiento de las enfermedades, el desarrollo de nuevos dispositivos médicos y la nutrición. Dentro de los llamados beneficios anticipados del PGH figuran la posibilidad de valorar la predisposición genética a las enfermedades, el diseño racional de nuevos medicamentos y la prescripción de estos ajustada a las peculiaridades genéticas de cada paciente (farmacogenómica).

Los proyectos de secuenciación masiva del ADN han contribuido al desarrollo de la bioinformática y al crecimiento exponencial de los bancos de información biológica que se almacenan en grandes bases de datos, la mayoría de las cuales son accesibles de forma gratuita a través del INTERNET. Sin la anotación facilitada por las herramientas bioinformáticas, la información desordenada del PGH sería de muy poca utilidad. Además de reunir y organizar toda esta información, es necesario analizar y descubrir en las secuencias de nucleótidos o en las aminoácidas que estas codifican, las semejanzas y las diferencias de cada hallazgo, contra lo conocido hasta entonces (Choudhuri, 2003).

1.2. Antecedentes.

1.2.1. Estado actual de los Proyectos de secuenciación del Genoma Humano y del Chimpancé.

La versión de la secuencia del genoma humano anunciada en el 2003 por el PGH, cubría el 92.3% de éste. Posteriormente se llevó a cabo una secuenciación adicional para cubrir los huecos, resolver las ambigüedades y reducir los errores a no más de uno por cada 10,000 bases (norma convenida para el consorcio de laboratorios que encabezan el PGH). La versión del genoma humano finalizada en mayo de 2006, proporciona un grado de confianza de la lectura de cada posición nucleotídica de nueve veces (Facts About Genome Sequencing, http://www.ornl.gov/sci/techresources/Human_Genome/faq/seqfacts.shtml#complete).

La culminación del PGH en términos técnicos significa que se obtuvo la secuencia de la molécula del ADN con una gran exactitud (99.99 por ciento) y cobertura (alrededor del 99 por ciento de las regiones del genoma humano que contienen genes). Únicamente el 1% restante de la cadena no pudo ser resuelta de manera segura con la tecnología actual, correspondiendo a regiones genómicas enriquecidas en secuencias altamente repetitivas,

comúnmente encontradas en los centrómeros. Secuenciar estas regiones restantes requiere proyectos de investigación individual (NHGRI, <http://www.genome.gov/11510908>).

El saldo del PGH es la secuencia completa de los 3,200 millones de nucleótidos que lo componen, el mapa que ubica a los cerca de 30,000 genes que ahí se albergan y la identificación de alrededor de 1,400 de estos que causan enfermedades monogénicas. Además, evidenció que los seres humanos comparten el 99.9% de la secuencia de sus genomas y que las variaciones más comunes son los polimorfismos de un solo nucleótido, conocidos como SNPs (del inglés, Single Nucleotide Polymorphism). Estas variaciones se ubican en posiciones separadas en promedio de 600 a 800 nucleótidos y hasta el momento se han identificado más de 3.2 millones de ellas. El número de las posibles combinaciones de SNPs hacen único a cada miembro de nuestra especie. Así, la individualidad genómica da lugar a la individualidad bioquímica, responsable de la predisposición a padecer enfermedades comunes, diferencias en la respuesta a medicamentos y vacunas, o en la reacción a factores medioambientales (Instituto Nacional de Medicina Genómica, México. http://inmegen.gob.mx/index.php?option=com_content&task=view&id=88&Itemid=213&mo=1&yr=2007&lang=es).

Investigaciones previas han identificado más de cien regiones del genoma humano que contienen variaciones genéticas asociadas con el riesgo de enfermedades comunes, como la diabetes, la enfermedad de la arteria coronaria, los cánceres de mama y de próstata, la artritis reumatoide, la enfermedad intestinal inflamatoria y la degeneración macular asociada con la edad. Sin embargo, como el mapa del genoma actual no es muy detallado en este punto, a menudo es necesario secuenciar para identificar las variantes genéticas precisas que causan una enfermedad.

Con el propósito de crear el mapa humano de variación genética más detallado hasta la fecha y médicamente útil, este año se ha conformado un consorcio internacional de investigación que ha anunciado el Proyecto de 1000 genomas, el cual involucrará la secuenciación del genoma de por lo menos mil personas de todo el mundo. Este proyecto recibirá el principal apoyo del Instituto Sanger en Hinxton (Inglaterra), el Instituto de

Genómica Shenzhen en Beijing (China) y el Instituto Nacional de Investigación del Genoma Humano en Maryland (Estados Unidos).

Utilizando la experticia de equipos de investigación multidisciplinarios, el Proyecto de 1000 genomas desarrollará un nuevo mapa del genoma humano que proporcionará una importante visión de las variaciones de ADN biomédicamente relevantes. Los datos generados a partir de este ambicioso proyecto se pondrán rápidamente a disposición de la comunidad científica mundial, a través de las bases de datos públicas. Este nuevo mapa ayudará a los científicos a detectar más rápidamente variantes genéticas asociadas con enfermedades y a promover esfuerzos para usar la información genética con el fin de desarrollar nuevas formas de diagnosticar, tratar y prevenir enfermedades con componente genético (GOLD, 2008. <http://www.1000genomes.org/files/1000Genomes-NewsRelease.pdf>).

La infraestructura heredada y el entusiasmo generado por el PGH, además de dar inicio a varias iniciativas, como el referido Proyecto de los 1,000 genomas, ha dado un nuevo impulso a la genómica comparativa, cuyo propósito es entender a través del análisis y la comparación de genomas de diferentes especies, cómo las especies han evolucionado y cuál es la función de los genes y las regiones no codificantes del genoma. Esto incluyó en la lista de las especies seleccionadas para secuenciar sus genomas, al chimpancé (*Pan troglodytes*), la especie más cercana evolutivamente a los humanos.

Como resultado de los avances del Proyecto del Genoma del Chimpancé, en diciembre de 2003 se obtuvo el primer borrador de la secuencia de este primate, con una cobertura de 4X (cuatro lecturas completas del genoma). El Consorcio de secuenciación y análisis del chimpancé implementó una secuenciación adicional que hizo posible que publicara en marzo de 2006 un ensamblaje mejorado de este borrador genómico con una cobertura de 6X.

Si bien es cierto que el borrador de la secuencia genómica del chimpancé reveló que el genoma de éste difiere en un porcentaje muy bajo con respecto al del humano, aún hay secuencias que faltan y errores típicos que incluyen problemas de ensamblaje de secuencias

repetidas y traslapes no determinados (asociados a polimorfismos) que originan duplicaciones artificiales. El genoma del chimpancé es una pieza más en la lista de los genomas de vertebrados secuenciados, que seguro será el de mayor utilidad para comprender la biología y evolución humanas.

1.2.2. Genómica comparativa como herramienta de estudio.

Con la secuencia de un genoma se obtiene más que largas filas de nucleótidos. En ella está contenida la información a la que se desea acceder, pero se necesita conseguir una manera de interpretarla. El análisis de genomas se refiere a la tarea de entender qué dice la secuencia de un genoma; básicamente qué genes contiene, dónde se encuentran y qué funciones realizan las proteínas que codifican (Rouze *et al.*, 1999; Stein, 2001; Rust *et al.*, 2002). A lo largo de la evolución, los cambios que sufren los genomas y sus genes, están sometidos a la selección natural. De las ventajas o inconvenientes que estos cambios acarreen, dependerá que se consoliden en las siguientes generaciones. La comparación de genomas, genes y proteínas de distintas especies es la aproximación más intuitiva y directa para interpretar “el manual de instrucciones” que hay en los genomas (Koonin, 2001; Salzberg, 2003).

Cuando se compara la secuencia de aminoácidos de proteínas que realizan una misma función en organismos distintos, se observa que son parecidas y que se puede establecer una correspondencia entre los residuos aminoacídicos en ciertas posiciones de unas y otras, existiendo algunas que son idénticas y otras donde hay diferencias. En el contexto de la evolución molecular, estas proteínas se parecen porque tienen un origen común y las diferencias se deben a que a lo largo del tiempo los genes han divergido por la acumulación de cambios o mutaciones en sus secuencias (Alberts 2006).

La comparación de proteínas ortólogas (aquéllas que tienen un origen común) proporciona mucha información acerca de cómo se ha producido la evolución de los organismos y las proteínas mismas, ya que el hecho de que unas posiciones aminoacídicas se conserven, está relacionado con cómo son clave para la estructura o la función de la proteína (Miller *et al.*, 2004), así como también con la importancia que tenga esta última para la sobrevivencia del

organismo. En este sentido, los residuos más directamente implicados en el desempeño de la función, suelen presentar menor variación, ya que cambios en estos pueden provocar una pérdida de función (Alberts 2006; Khaitovich *et al.*, 2006).

La importancia de estos estudios comparativos de proteínas homólogas fue destacada por primera vez hace medio siglo (Zuckerkandl y Pauling, 1965) y desde entonces ha sido puesta de manifiesto una y otra vez en numerosos trabajos. Pero no siempre es fácil obtener información a partir de la comparación de secuencias, ya que muchas veces lo que observamos no está ahí por ser óptimo, sino como una mera consecuencia de un derrotero de la historia evolutiva; como dijo en público el doctor Sydney Brenner, ganador del premio Nobel de medicina: "Si las matemáticas son el arte de lo perfecto y la física es el arte de lo óptimo, la biología no es más que el arte de lo satisfactorio: cualquier cosa sirve, siempre que funcione". Esto se puede expresar de forma más precisa utilizando los términos de selección positiva, negativa y neutra, que se refieren a cómo resulta un cambio para un organismo (Abascal, 2003).

Las proteínas con secuencias suficientemente parecidas suelen tener un origen común (Zuckerkandl y Pauling, 1965) y presentar una arquitectura tridimensional similar (Chothia y Lesk, 1986), pero frecuentemente realizan funciones distintas (Henikoff *et al.*, 1997), aunque también pueden traslaparse en otras. Por función de una proteína se entiende el papel que ésta desarrolla, como puede ser la participación en una ruta metabólica, la transmisión de señales al interior de la célula y la regulación de la función de genes, entre otras.

La duplicación génica está considerada como uno de los fenómenos más relevantes en la evolución de las proteínas y de sus nuevas funciones (Ohno *et al.*, 1968; Watts y Watts, 1968; Ohta, 1989; Lang *et al.*, 2000). Se ha postulado que al existir dos copias de un gen, la presión selectiva frente a cambios en residuos importantes es menor, ya que si uno de los genes pierde la función, aún queda la del otro. Esto facilita la exploración de nuevos nichos funcionales para las proteínas miembros de la familia, que empiezan a cambiar. Existen

familias de proteínas en las que este fenómeno es especialmente notorio, por lo que en estas se puede observar gran variedad de funciones entre sus miembros (Alberts 2006).

En muchas situaciones las familias de proteínas que tienen un origen común y una estructura tridimensional parecida, exhiben actividades enzimáticas similares. Por ello se podría considerar que estas proteínas realizan una misma función, pero el caso es que esta actividad puede contribuir a distintos procesos celulares, lo que pone de manifiesto uno de los principales problemas encontrados cuando se quiere anotar la función de las proteínas de forma automática. La interpretación de la función tiene un componente subjetivo muy importante: según la formación que tenga el científico que anota la proteína, prestará mayor atención a unos u otros aspectos de la función (Abascal, 2003).

Los investigadores han aprendido mucho sobre la evolución y función de los genes humanos examinando a sus contrapartes en modelos animales. Ello al comparar sus genomas para determinar la similitud de las secuencias, localización de los genes, la longitud y número de exones en éstos, las regiones genómicas altamente conservadas y la cantidad de ADN no codificante (Birney *et al.*, 2007). La secuenciación del genoma del chimpancé proporciona un medio para estudiar aspectos estructurales y funcionales de la evolución del genoma humano, desde su divergencia a partir de un antepasado común hace aproximadamente entre 5 y 7 millones de años (Leakey *et al.*, 1998; Glazko y Nei, 2003; Kumar *et al.*, 2005).

1.2.3. Comparaciones genómicas humano-chimpancé.

Los estudios de ambos genomas muestran patrones de evolución altamente correlacionados en los genes que codifican proteínas. Al alinearse los genomas del humano y del chimpancé, se estima que éstos difieren en 1.23% de sus posiciones nucleotídicas, abarcando aproximadamente 35 millones de mutaciones puntuales, que incluyen aproximadamente cinco millones de inserciones, deleciones, duplicaciones e inversiones (Mikkelsen *et al.*, 2005; Khaitovich *et al.*, 2006).

Teniendo en cuenta la pequeña diferencia entre las secuencias del humano y chimpancé, se puede presumir que en ella están contenidos los cambios moleculares cruciales que dirigieron el proceso evolutivo para hacer posible que emergiera nuestra especie (Khaitovich *et al.*, 2006; Kehrer-Sawatzki y Cooper, 2007).

La mayor parte de las divergencias fenotípicas es resultado de cambios genómicos específicos que han dado lugar a diferencias en la expresión genética, así como a ganancias, pérdidas y cambios en la función proteica (King y Wilson, 1975; revisado en Carroll, 2003; Olson y Varki, 2003; Enard y Pääbo, 2004; Carroll, 2005; Goodman *et al.*, 2005; Varki y Altheide, 2005; Khaitovich *et al.*, 2006).

El reciente progreso en las comparaciones de los genomas de humano y chimpancé ha revelado que las inserciones y deleciones han alterado considerablemente el paisaje genómico de estas especies después de la separación de ambos linajes (Khaitovich *et al.*, 2006; Kehrer-Sawatzki y Cooper, 2007). La secuenciación y comparación de ambos genomas han comenzado a revelar cambios genéticos de posible peso en la divergencia entre ambas especies. Sin embargo, la aún baja resolución del ensamblaje del genoma del chimpancé desafía la habilidad de los métodos actuales para identificar muchos de estos cambios estructurales (Kehrer-Sawatzki y Cooper, 2007). Por ende se requiere de investigaciones de genómica comparativa de regiones génicas específicas, de las que mientras tanto se puedan derivar lecciones.

1.2.4. Genómica comparativa de *loci* definidos: el proyecto ENCODE.

El genoma humano todavía alberga muchas incógnitas. Su secuenciación en el 2003 fue un gran logro, pero tan sólo el primer gran paso. El mapa está listo, pero todavía falta mucho por tener del manual de instrucciones completo. Seguro que por el camino habrá que deshacerse de algunas de las ideas más arraigadas hasta el momento sobre su estructura y funcionamiento. Así lo acaban de demostrar los resultados del proyecto internacional ENCODE o gran Enciclopedia de los Elementos del ADN (Birney *et al.*, 2007), que desarrolló un consorcio internacional de investigadores encabezados por el Instituto

Nacional de Investigación del Genoma Humano (NHGRI), el mismo instituto estadounidense que dirigió el PGH.

Hasta hace poco tiempo se creía que los genes eran unidades independientes, separadas por secuencias de ADN inservibles y sin ninguna función. También se pensaba que los únicos que podían codificar eran los genes y, en concreto, que cada uno codificaba una sola proteína. Pues bien, los resultados del proyecto ENCODE han revelado que los genes no son unidades independientes, sino que muchos se superponen y actúan en red compartiendo información. Además, es muy frecuente que codifiquen más de una proteína y ese ADN aparentemente inútil que ocupa aproximadamente el 95% del genoma resulta que importa y mucho. Por si fuera poco, todos estos elementos interactúan de manera compleja y sobrepuesta (Birney *et al.*, 2007). Así pues, la estructura y el funcionamiento de un sólo gen presenta todavía muchas incógnitas. Con toda esta información sobre la mesa, resulta imperativo replantear algunas ideas acerca de los genes y la evolución de los elementos funcionales.

En su fase piloto, el proyecto ENCODE se centra en el estudio de 44 regiones que cubren el 1% del genoma humano, de las cuales, 30 han sido seleccionadas al azar y 14 corresponden a genes o grupos de genes ya identificados. Un aspecto clave para la selección de las regiones incluidas en este estudio ha sido la disponibilidad de datos de secuencias ortólogas para un amplio número de especies de mamíferos que hicieran posible el desarrollo de estudios de genómica comparativa (Consorcio del Proyecto ENCODE, 2004).

Esta fase piloto del proyecto sirve para probar y evaluar la efectividad y eficacia de diversos métodos y tecnologías que posteriormente se aplicarán a todo el análisis del genoma. El proyecto ENCODE se propone crear una enciclopedia total de los elementos funcionales codificados en la secuencia del ADN, catalogando su identidad y ubicación precisa dentro del genoma; esto mediante el empleo de tecnologías de alto rendimiento y herramientas bioinformáticas (Consorcio del Proyecto ENCODE, 2004; Birney, *et al.*, 2007).

Algunos de los primeros hallazgos incluyen importantes descubrimientos sobre el papel de las regiones del ADN que no participan en la codificación de proteínas. Los nuevos datos muestran que sólo una mínima parte de la secuencia no cumple una función biológica. El porcentaje del genoma considerado hasta ahora como “ADN basura”, tiene en realidad papeles reguladores esenciales (http://www.cbse.ucsc.edu/news/2007/06/13/encodeposter_nature0607sm.jpg).

Cabe mencionar que entre las familias génicas seleccionadas en la fase piloto del Proyecto ENCODE no se encuentra incluida la de la hormona del crecimiento (GH), modelo de estudio de nuestro laboratorio. En el presente trabajo enfocamos nuestros análisis al *locus* GH en primates, grupo en el que experimentó una rápida evolución, caracterizada por pasar de un *locus* uni-gen a uno multi-génico, seguido por la divergencia de algunos de sus miembros. Nuestro esfuerzo está dirigido a reconstruir la secuencia de eventos que produjeron esta dramática transformación.

1.2.5. Evolución del *locus* GH en primates.

La GH es una hormona que se sintetiza y secreta en la glándula pituitaria de todos los mamíferos para regular el crecimiento y varios aspectos centrales del metabolismo. Está presente en los vertebrados y en la mayoría de los mamíferos su *locus* posee un sólo gen. Sin embargo, en primates pasó a ser un *locus* multigénico (Wallis y Wallis, 2006).

Los primates surgen en el período Terciario de la era Cenozoica y se clasifican en dos subórdenes: los prosimios y los antropoides. En el oligoceno, evolucionaron los antropoides a partir de un ancestro común y se dividieron en tres grupos: *Cebioidea* o monos del nuevo mundo (MNM), *Cercopithicoidea* o monos del viejo mundo (MVM) y *Hominoidea* que incluye a los grandes monos (GM) y al hombre (figura 1).

En los prosimios la GH es codificada por un solo gen, cuya secuencia es muy similar a la ancestral encontrada en los mamíferos. Sin embargo, en los antropoides se han descubierto *loci* GH integrados por múltiples genes no alélicos relacionados (Chen *et al.*, 1989; Wallis y Wallis, 2002; Revol De Mendoza *et al.*, 2004; González-Álvarez *et al.*, 2006). En

general, las GHs en este grupo de primates presentan gran similitud en sus secuencias y en sus formas maduras constan de 191 aminoácidos.

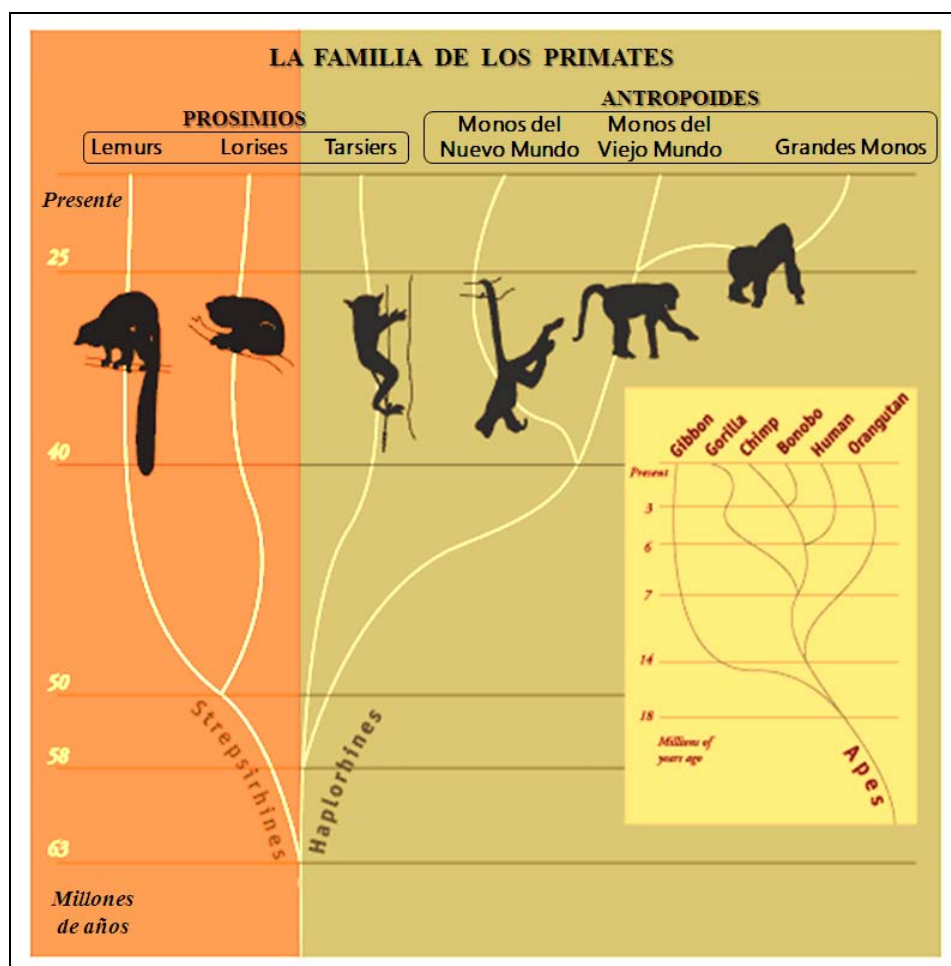


Figura 1. Evolución de los primates. La construcción de este árbol filogenético se apoyó en registros paleontológicos. Modificado de Great Ape Trust, <http://www.greatapetrust.org/primates/index.php>.

Después de la separación de prosimios y antropoides, pero antes de la divergencia de los MNM, los MVM y los homínidos, múltiples eventos de duplicaciones génicas condujeron a un aumento del número de genes GH que evidencian la notable expansión que ha sufrido este *locus* en primates (Revol De Mendoza *et al.*, 2004; González-Álvarez *et al.*, 2006; Wallis y Wallis, 2006).

La multiplicación génica del *locus* GH y la divergencia de sus miembros fueron acompañadas por cambios que resultaron en diferencias en cuanto a la expresión en los

tejidos de los genes miembros del *locus*, con al menos dos especificidades: hipófisis y placenta. Además del gen del tipo GH cuya expresión siguió restringida a la hipófisis, aparecieron genes de hormonas somatomamotropinas coriónicas (CSH) e incluso una variante de la GH que tienen a la placenta como su sitio de síntesis (Chen *et al.*, 1989; Golos *et al.*, 1993; Revol *et al.*, 2004, Ye *et al.*, 2005).

Hasta el momento no está claro si algunos genes GH son expresados en placenta en MNM. Sin embargo, a partir de MVM los miembros de este agrupamiento acumularon sustituciones aminoacídicas que predicen la adquisición de nuevas propiedades distintivas de los miembros placentarios del *locus* humano, permaneciendo como un enigma el origen de las diferencias en especificidad tisular de la expresión desplegada por los genes del *locus*. Las secuencias de múltiples genes relacionados a GH se han descrito y analizado en antropoides (Revol *et al.*, 2004; Ye *et al.*, 2005; Li *et al.*, 2005; Wallis y Wallis 2006), pero a la fecha, la familia GH solo ha sido estudiada a detalle en el humano.

1.2.6. Anatomía y fisiología del *locus* HGH.

1.2.6.1. Estructura del *locus* HGH.

El *locus* GH en el humano se localiza en la banda q22.4 del brazo largo del cromosoma 17 (George *et al.*, 1981; Harper *et al.*, 1982), abarca una extensión de 66.5 kpb y contiene cinco genes estructuralmente conservados que se encuentran orientados en la misma dirección transcripcional (figura 2) (Chen *et al.*, 1989). El *locus* hGH se encuentra flanqueado en su extremo 5' por el gen CD79b específico de los linfocitos B (Jae *et al.*, 2006) y en su extremo 3' por el gen TCAM específico de testículos (Sakatani *et al.*, 2000; Atsushi *et al.*, 2004).

En la región 5' del *locus* se ubica el gen hGH-N (referido como normal), cuya expresión es específica de los somatotrofos de la glándula pituitaria anterior. Los restantes cuatro genes se expresan en el sincitiotrofoblasto de la placenta, dos de los cuales (hCSH-A y hCSH-B) codifican para una hormona madura idéntica (Barrera-Saldaña *et al.*, 1983), mientras que el restante, hCSH-L, es considerado un pseudogen por una mutación en su segundo intrón que evita el procesamiento correcto de sus transcritos (Resendez *et al.*, 1990). El gen

placentario restante es el hGH-V, el cual codifica para una variante de la GH que se postula sustituye en su función a la GH hipofisiaria durante el embarazo (Chen *et al.*, 1989), (Alsat *et al.*, 1997).

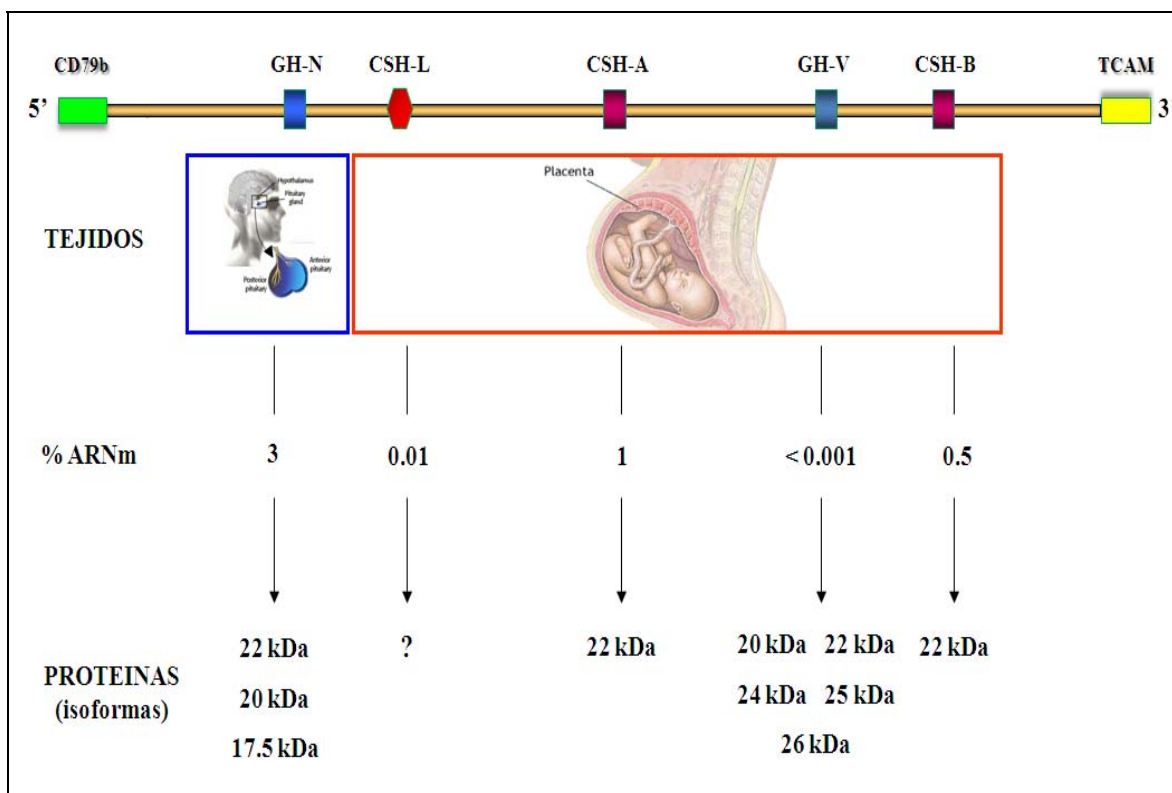


Figura 2. Anatomía del locus hGH. Se muestra la organización de los genes dentro del locus, el tejido donde se expresan, el porcentaje del ARNm en dicho tejido y las isoformas proteicas que generan.

Las unidades transcripcionales de estos cinco genes no ocupan más de 2 kpb cada una y están separadas por cuatro regiones intergénicas: dos de 13 kpb, una de 6.5 kpb y una de 6 kpb (Chen *et al.*, 1989). En estas regiones se han identificado tanto secuencias repetitivas, como elementos que participan en la regulación de la expresión de los genes. De hecho, dentro del locus se encuentran 48 copias de secuencias repetitivas de la familia *Alu* y una de la familia *Kpn* (figura 3). Las inserciones *Alu* ocurren en ambas orientaciones y tres de ellas están truncadas. Cabe señalar que más del 20% de las secuencias nucleotídicas del locus hGH corresponde a estos elementos repetitivos, cuyo patrón de distribución, semejante para los cinco genes, sugiere su posible implicación en la ruta evolutiva que dio forma al locus (Chen *et al.*, 1989).

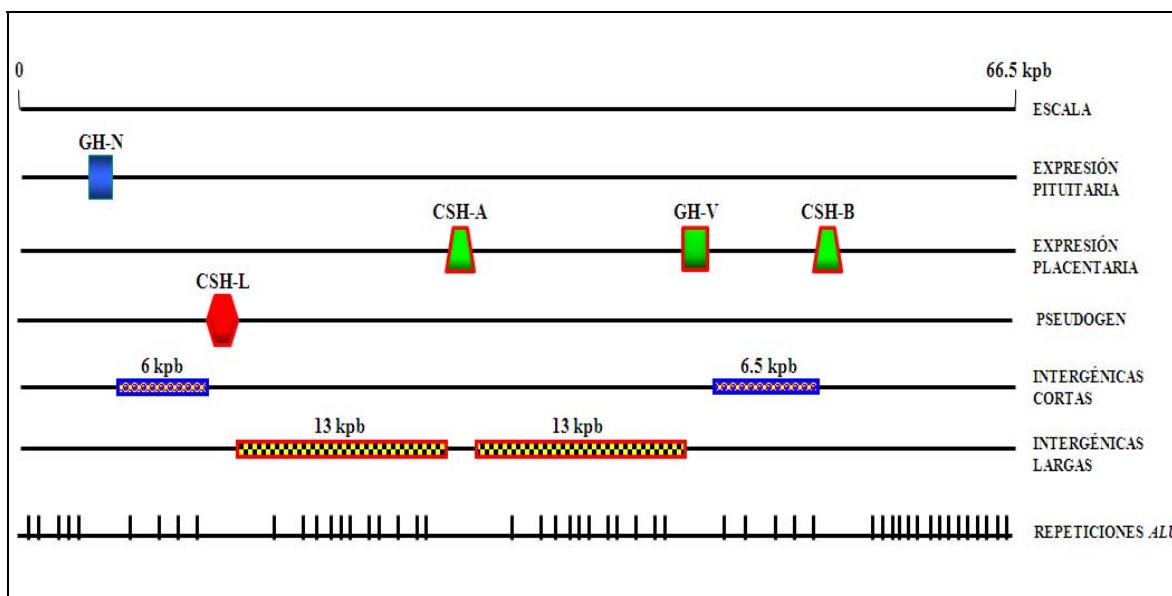


Figura 3. Distribución de elementos *Alu* en el locus *hGH*. Los genes están separados por regiones intergénicas cuyas longitudes se señalan. Las secuencias *Alu* están distribuidas en las regiones que flanquean a los genes, representadas aquí por líneas verticales.

A pesar de sus diferencias en cuanto a especificidad tisular, niveles de expresión y funciones de sus proteínas, los genes GH/CSH presentan una elevada similitud en sus secuencias nucleotídicas (superior al 90%) y exhiben la misma anatomía: cinco exones separados por cuatro intrones. El análisis estructural de las proteínas globulares que codifican ha revelado que poseen cuatro hélices alfa y contienen dos puentes disulfuro. Recién sintetizadas estas pre-hormonas presentan un péptido señal de 26 aminoácidos que dirige su transporte a través de la membrana del retículo endoplasmático rugoso hacia el aparato de Golgi, desde donde son secretadas en su forma madura al torrente sanguíneo (Barrera *et al.*, 1982; Handwerger *et al.*, 2000).

El gen hGH-N genera a través de una remoción diferencial de su segundo intrón durante la maduración del transcrito (“splicing” alternativo) dos isoformas: una de 22 kDa que es la más conocida y abundante (90%) y otra de 20 kDa. A partir del gen hGH-V se derivan también varias isoformas, siendo el ARNm más abundante de este gen en la placenta al final del embarazo el que codifica también para una proteína de 22 kDa (Frankenne *et al.*, 1987).

1.2.6.2. Papel funcional de las hormonas del *locus* HGH.

Resulta más fácil comprender el sentido de los eventos evolutivos que originaron el *locus* hGH al abordar el papel que juegan las proteínas codificadas por sus miembros en la fisiología humana. La principal función de la hGH consiste en regular el crecimiento lineal postnatal y diversas actividades metabólicas. Esta hormona interviene en procesos como la homeostasis de la glucosa, la síntesis proteica en el tejido muscular, la lipólisis en el tejido adiposo, la reparación tisular, la regeneración ósea (Farreras *et al.*, 2000), la neurogénesis y neuroprotección, el aprendizaje y la memoria (Sonntag *et al.*, 2005).

La expresión de GH-V se restringe al sincitiotrofoblasto de la placenta (Coocke *et al.*, 1988; 1991) y sus niveles se asocian con los índices de crecimiento intrauterino, peso al nacer y con los niveles fetales del factor de crecimiento parecido a la insulina I (IGF-I) (Verhaeghe *et al.*, 2002; Chellakooty *et al.*, 2004; Reygård *et al.*, 1998). Esta hormona presenta actividad somatogénica (MacLeod *et al.*, 1991), lactogénica (Níkel *et al.*, 1990) y como se señaló arriba, sustituye a la GH-N durante la segunda mitad de la gestación (Frankenne *et al.*, 1987). Además se le ha implicado en el aumento de los niveles de IGF-I en el suero materno (Caufriez *et al.*, 1990).

Un papel similar se le adjudica a las hCSH, cuyos niveles se han asociado con el buen funcionamiento placentario y con los niveles fetales de IGF-1 (Handwerger *et al.*, 2000; Karabulut *et al.*, 2001). Esta hormona de 22 kDa se piensa que juega un importante papel como antagonista de la insulina en la regulación del metabolismo materno durante el embarazo y además de causar retención de nitrógeno, produce intolerancia a la glucosa (Josimovich *et al.*, 1964; McGarry *et al.*, 1972).

Los efectos contrarios a la insulina que tiene la hCSH comprenden la inhibición de la síntesis de glucógeno, la glucólisis, la lipogénesis y del ingreso de glucosa a las células, además la activación de la glucogenólisis, la proteólisis en el músculo y la gluconeogénesis a partir de los aminoácidos. Estas acciones tienen como propósito aumentar la concentración plasmática de glucosa, preservándola para su consumo por órganos como el cerebro materno y suministrar las cantidades adecuadas de glucosa y de nutrientes al feto

(<http://encolombia.com/obstetricia50399editorial.htm>). Además ejerce una actividad somatotrópica directa sobre los tejidos fetales. También puede acelerar la lipólisis (Felber *et al.*, 1972; Williams *et al.*, 1978), por lo que se piensa que puede servir para liberar ácidos grasos durante la gestación, para asegurar el suministro de energía al feto (Walker *et al.*, 1991; Kim *et al.*, 1971). Durante el tercer trimestre del embarazo se produce un incremento en la síntesis de esta hormona, llegando a representar el 10% de la producción total de las proteínas placentarias (Jacquemin *et al.*, 1994).

1.2.7. Regulación de la expresión de los genes del *locus* GH en el humano.

La expresión génica está estrictamente controlada por mecanismos que operan a niveles de organización nuclear, estructura cromosomal y modificaciones del ADN. Cuando la cromatina es organizada en una configuración altamente compactada, el ADN es inaccesible para la mayoría de los activadores transcripcionales, dando como resultado una represión génica. La activación génica implica una descondensación controlada de la cromatina, acompañada de modificaciones covalentes (acetilación) de las histonas (Kimura *et al.*, 2004).

La activación de un *locus* como el GH también puede influenciarse por el efecto de una región de control del *locus* (LCR) (Kimura *et al.*, 2004). Este tipo de elemento tiene la capacidad de activar la expresión de genes blanco y de establecer, en la cromatina, dominios autónomos de regulación transcripcional. Además se caracterizan por la presencia de sitios hipersensibles a la enzima DNAsa I (Hipersensitive Sites, HS). Esta hipersensibilidad está asociada con cambios en la estructura de la cromatina, que a su vez indican eventos de actividad transcripcional de genes blanco (Shewchuk *et al.*, 2002; Elefant *et al.*, 2000). En general, una estructura compacta de la cromatina se asocia con una ausencia de actividad transcripcional y la unión de las histonas al ADN favorece esta estructura particular de la cromatina. El proceso de activación transcripcional se asocia con el cambio a una estructura abierta de la cromatina que se logra mediante la acetilación de las histonas.

La LCR del *locus* hGH está marcada por un conjunto de sitios hipersensibles a DNasa I (HSI-HSV) localizados entre 15 y 32 kpb antes del *locus* (figura 4). Los sitios HSI y HSII, son específicos para la expresión en la pituitaria y la acetilación de las histonas en una región centrada en estos sitios se asocia con la expresión del gen hGH-N en esta glándula. El sitio HSIV es específico de placenta y la acetilación de histonas asociada a la expresión de los genes placentarios está limitada a la región que abarca HSIII-HSV. Los sitios HSIII y HSV, localizados a 28 kpb y 32 kpb río arriba del gen hGH-N, se encuentra en ambos tipos celulares.

La expresión diferencial de los miembros del agrupamiento génico del *locus* hGH en placenta e hipófisis depende en gran medida de un control ejercido por los distintos subgrupos de la LCR. (Jones *et al.*, 1995; Elefant *et al.*, 2000, Kimura *et al.*, 2004). Regiones internas del *locus* multigénico son selectivamente acetiladas en la cromatina placentaria y una de estas regiones cromatínicas comprende precisamente al elemento P de los genes placentarios, situado enfrente de estos e implicado en la represión de su expresión en hipófisis (Kimura *et al.*, 2004).

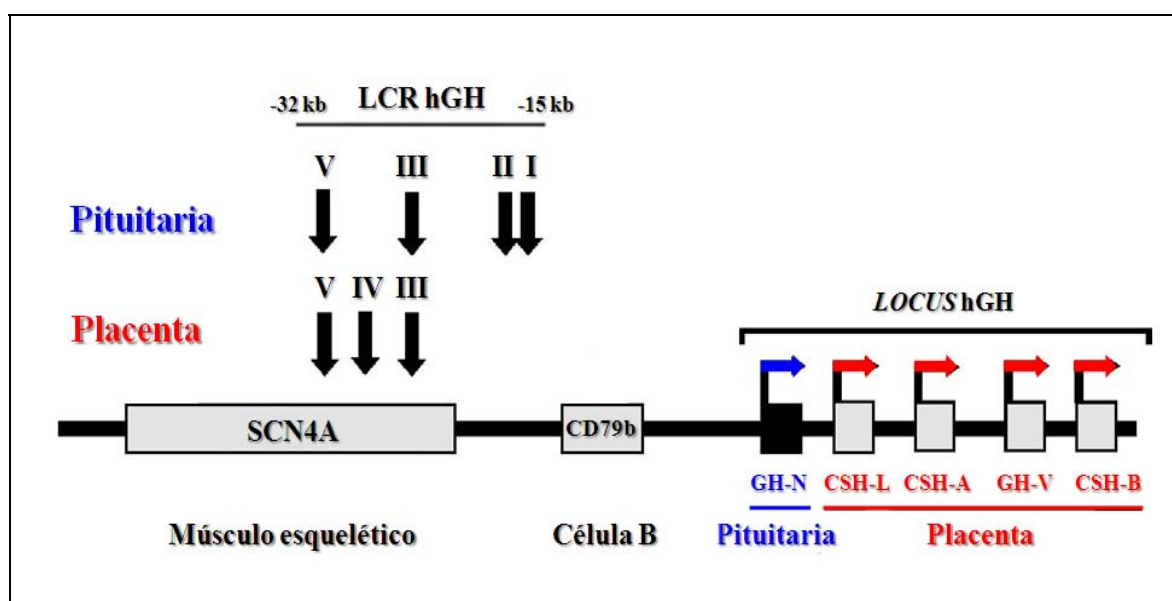


Figura 4. Región de control del *locus* hGH. Se muestra el gen hGH-N (cuadrado negro) y los cuatro genes de expresión placentaria (cuadrados grises) que conforman al *locus* hGH. Con rectángulos grises se representan los genes no relacionados, pero estrechamente ligados SCN4A y CD79b. Las flechas horizontales muestran la orientación transcripcional de los genes del *locus* hGH. La posición de cada sitio HS en la LCR del *locus* hGH está marcado con flechas verticales. Modificado de Shewchuck *et al.*, 2006.

Uno de los puntos claves de la regulación de la expresión de los genes es el control transcripcional (figura 5). En esta categoría está incluida la acción de los promotores, la presencia de secuencias regulatorias potenciadoras (enhancers) y la interacción entre múltiples proteínas activadoras o inhibidoras que actúan mediante su unión a secuencias específicas de reconocimiento al ADN. El control transcripcional no sólo se ocupa de "encender" o "apagar" genes, sino también de regular la intensidad de la transcripción de los genes encendidos.

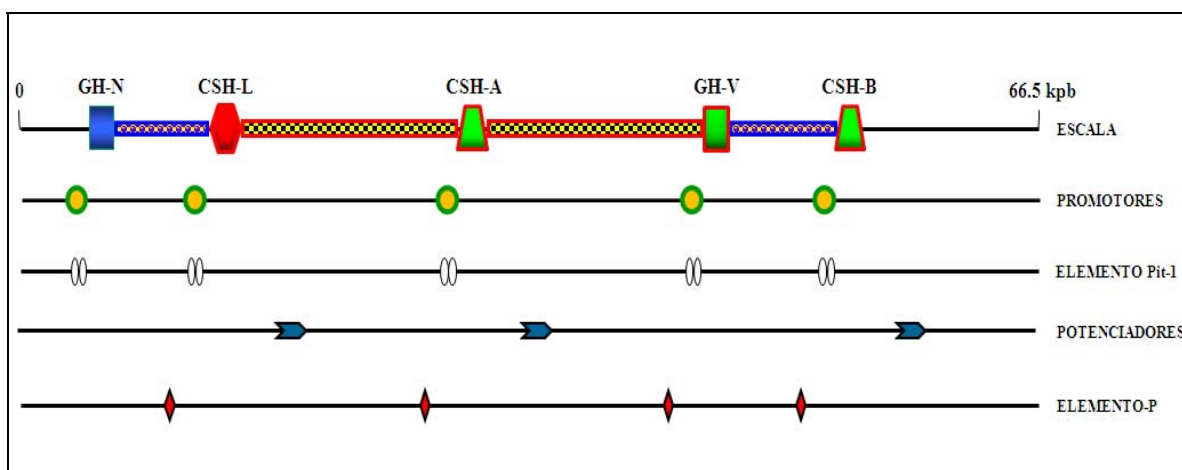


Figura 5. Localización de sitios que participan en la regulación de la expresión de los genes del locus hGH. Con círculos se muestra la ubicación de las regiones promotoras de los genes del locus y con dos óvalos se representan los elementos Pit-1 presentes en estos. Con punta de flecha se representa el potenciador placentario después de cada gen CSH, con rombos se indica el elemento P o inhibidor antes de cada gen de expresión placentaria.

La expresión del gen hGH-N en los somatotrofos de la pituitaria anterior es dirigida por el factor transcripcional Pit-1, también conocido como GHF1 (Karin *et al.*, 1990). Por ser Pit-1 un factor específico de tejido hipofisiario (Ingraham *et al.*, 1988), aparentemente no está involucrado en la inducción de la expresión de los genes placentarios. Este factor presenta dos dominios conservados: un dominio POU importante para la activación transcripcional y un dominio homeobox (HB) que intervienen en su unión al ADN, ambos ubicados en el extremo carboxilo terminal de la proteína (Schanke *et al.*, 1997).

El factor Pit-1 se une al promotor del gen GH-N en dos sitios: un sitio proximal (-92 a -65) y un sitio distal (-130 a -105) (Courtois *et al.*, 1992). A pesar de que los promotores de los genes placentarios también presentan sitios de unión a Pit-1, estos están sujetos a una

represión transcripcional mediada por una región conservada de 263 nucleótidos denominada elemento P o inhibidor hipofisiario, situada aproximadamente a 2 kpb enfrente del sitio de inicio de la transcripción (Norquay *et al.*, 2001), a la cual se unen en la hipófisis proteínas que luego interfieren con la unión de Pit-1 a los promotores de dichos genes, impidiendo su expresión en esta glándula (Nachtigal *et al.*, 1993; Barrera-Saldaña *et al.*, 1998). El elemento P contiene dos dominios: PSE-A y PSE-B (figura 6). El primero interacciona con miembros de la familia del Factor Regulador X (RFX), mientras que el segundo lo hace con el Factor Nuclear-1 (NF-1). (Norquay *et al.*, 2001; Nachtigal *et al.*, 1993; Norquay *et al.*, 2003). Un grupo de factores *trans* comunes - Sp1, USF/MLTF y NF-I - están involucrados en el establecimiento del nivel basal de la transcripción del gen hGH-N en la hipófisis (Jacquemin *et al.*, 1994).

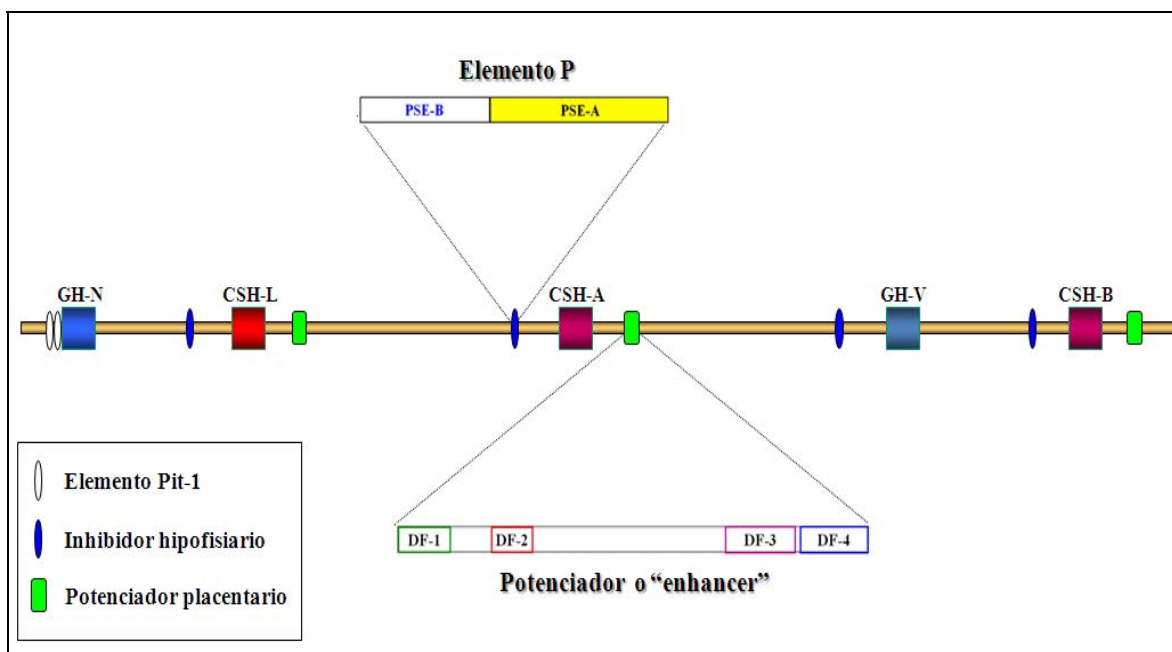


Figura 6. Detalles de los sitios de regulación de la expresión de los genes del locus hGH. Los dos sitios Pit-1 presentes en el promotor de hGH-N se esquematizan antes de dicho gen. El sitio del inhibidor hipofisiario o elemento P se representa antes de cada gen placentario, ampliados se destacan sus dominios PSE-A y PSE-B. El potenciador placentario se representa después de los genes CSHs, ampliados se resaltan los cuatro sitios de unión a proteínas nucleares de placenta que éste presenta: DF-1, DF-2, DF-3 y DF-4.

Aunque aún quedan muchas interrogantes sobre la regulación transcripcional en la placenta de los genes GH/CSH placentarios, ya se han realizado avances en este sentido. Para los genes hCSHs se ha identificado otro elemento regulador referido como elemento Iniciador o InrE, el cual se localiza entre la caja TATA y el sitio de iniciación de la transcripción.

Mutaciones en esta secuencia reducen en un 60% la expresión del promotor correspondiente (Jiang *et al.*, 1995).

Aproximadamente a 2.3 kpb después de los genes hCSH se localiza un potenciador, que incrementa la expresión de estos en la placenta (Rogers *et al.*, 1986; Jacquemin *et al.*, 1994, Barrera-Saldaña *et al.*, 1998). Ensayos de protección contra digestión con DNasa I han revelado, en este potenciador, la presencia de cuatro sitios de unión, denominados DF-1, DF-2, DF3 y DF-4, para proteínas nucleares de placenta (figura 6) (Jacquemin *et al.*, 1994), habiéndose demostrado que la actividad potenciadora del enhancer en placenta está mediada por un sinergismo entre los sitios DF-3 y DF-4 (Jacquemin *et al.*, 1996).

Se ha descrito que la función de este potenciador es independiente de su orientación y posición relativa (Rogers *et al.*, 1986), así como que su actividad requiere de la unión cooperativa a varios sitios localizados dentro de este elemento, del factor potenciador de la transcripción-1 (TEF-1) y de un factor potenciador específico de la placenta denominado CSEF-1. Los potenciadores de los tres genes hCSHs presentan una alta similitud en su secuencia, sin embargo, las copias en genes CSHL y CSHA muestran una débil actividad de estimulación transcripcional en comparación a CSHB, debido a mutaciones en sus respectivos sitios de unión a proteínas DF-3 (Jacquemin *et al.*, 1996).

1.2.8. Estado del *locus* GH del chimpancé al comenzar este trabajo.

Antes de comenzar el presente trabajo se conocía en detalle la secuencia y la organización del *locus* GH en el humano (Chen *et al.*, 1989), mientras que para el chimpancé se habían obtenido en nuestro laboratorio las secuencias completas de dos genes GHs, tres CSHs y la secuencia parcial de otro CSH. El hallazgo de estos seis genes discrepaba con las cuatro regiones intergénicas que se habían logrado amplificar para esta especie representativa de los grandes monos (Revol *et al.*, 2004). Por otro lado, como resultado de los avances del Proyecto Genómico del Chimpancé, también al momento de iniciar este trabajo se había recién liberado el primer borrador de la secuencia de su genoma. Con la escasa información disponible y sólo con una suposición sobre la organización del *locus* GH del chimpancé, sustentada en las secuencias de los genes y la pobre representación de dicho *locus* en el

ensamblaje genómico liberado por el consorcio de secuenciación, resultaba imposible entender los aspectos estructurales y funcionales de la evolución de esta región del genoma de tan cercano pariente del humano, que desde su divergencia a partir de un antepasado común hace aproximadamente unos 5 a 7 millones de años, han contribuido sin embargo a diferenciarlos.

1.3. Justificación del trabajo.

Para inferir los eventos que dieron lugar a la familia GH del genoma humano, resulta necesario recabar y ordenar la información del *locus* en el chimpancé, su pariente evolutivo más cercano. La comparación de ambos *loci* pudiera ilustrarnos la contribución de esta familia a la sobrevivencia y el éxito evolutivo de nuestra especie. El enfoque de genómica comparativa, aunado a las herramientas de biología molecular y celular, nos permitirá aspirar a inferir las fuerzas que moldearon la evolución de esta familia multigénica, generaron los mecanismos de su regulación, y subyacen en la relación estructura-función de las hormonas codificadas por sus miembros. Así mismo, a explotar las posibles aplicaciones biotecnológicas de este conocimiento y materiales.

CAPÍTULO II: OBJETIVOS

2.1. OBJETIVO GENERAL

Comparar los *loci* GH de chimpancé (*Pan troglodytes*) y humano (*Homo sapiens*), para identificar características distintivas de este último que pudieron haber contribuido a su éxito evolutivo.

2.2. OBJETIVOS ESPECÍFICOS

2.2.1. Desarrollar estrategias que permitan completar la secuenciación y el ensamblado del *locus* GH del chimpancé.

- a) Recabar y ordenar la información sobre el locus, disponible en las bases de datos de secuencias.
- b) Proponer arreglos de genes para describir la estructura del locus.

2.2.2. Determinar la organización y composición genómica del locus GH de chimpancé.

- a) Aislar y caracterizar los genes y las regiones intergénicas.
- b) Secuenciar los genes y las regiones intergénicas.
- c) Ensamblar y anotar dichas secuencias.
- d) Verificar el ensamblado del *locus* a través de la comparación de los resultados experimentales *versus* las predicciones en base a su secuencia.

2.2.3. Comparar entre los loci GH de humano y chimpancé sus:

- a) Genes (secuencias nucleotídicas y las inferidas de aminoácidos).
- b) Elementos involucrados en la regulación transcripcional.
- c) Patrones de elementos repetitivos presente en las regiones intergénicas.

2.2.4. Realizar inferencias sobre el posible rol de las características distintivas del locus humano en el éxito evolutivo de nuestra especie.

CAPÍTULO III: ESTRATEGIA GENERAL

La estrategia general (figura 7) para desarrollar el presente trabajo comprendió varias etapas. Primero se recabó y ordenó la información sobre el *locus* GH del chimpancé, disponible en las bases de datos de secuencias. Luego se extrajo DNA a partir de una cepa portadora de un BAC, cuyo inserto incluye al *locus* GH del chimpancé, a partir del cual empleando iniciadores consensos diseñados específicamente para este trabajo, se amplificaron los genes y las regiones intergénicas. En seguida los productos amplificados se caracterizaron con enzimas de restricción, clonaron y secuenciaron. Con las secuencias obtenidas y las recabadas de las bases de datos, se procedió a la reconstrucción y anotación de la región génica en estudio. Se concluyó realizando un análisis comparativo de este *locus* con su contraparte en el humano, teniendo en cuenta sus genes, las proteínas que éstos codifican, los elementos involucrados en la regulación transcripcional y el patrón de elementos repetitivos presente en las regiones intergénicas. Finalmente se plantearon hipótesis para tratar de explicar los principales hallazgos.

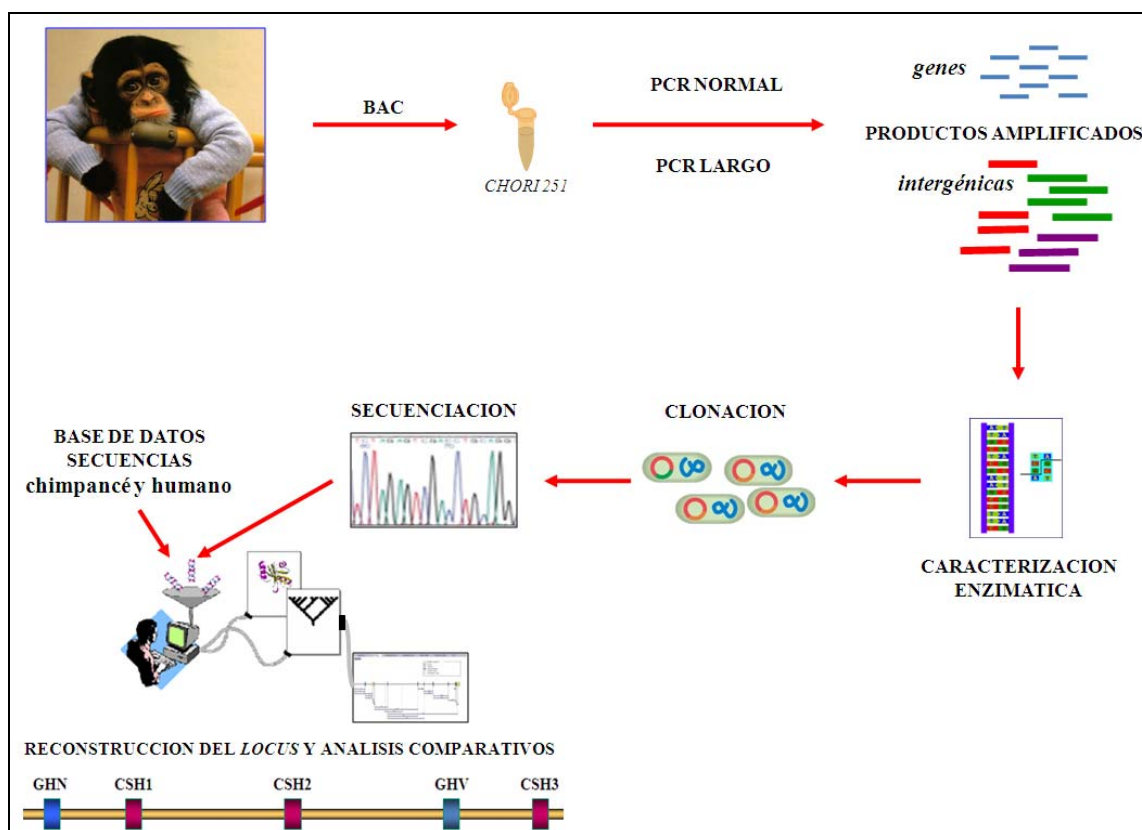


Figura 7. Estrategia general. Comprendió las etapas de extracción y purificación del DNA del BAC, amplificación de los genes y regiones intergénicas, caracterización enzimática, clonación, secuenciación y análisis bioinformático de las secuencias.

CAPÍTULO IV: MATERIAL Y MÉTODOS

4.1. Material.

4.1.1. Origen de los reactivos.

Los reactivos utilizados en este trabajo para la preparación de las soluciones, medios de cultivo y geles provinieron de las compañías Sigma Chemical Company (St, Louis, MO, EUA), Merck (Monterrey, México), Aldrich Chemical Company (Milwaukee, WI, EUA), United States Biochemical (Cleveland, Ohio, EUA) y DIFCO Laboratories (Detroit, MI, EUA).

El cloranfenicol presente en el medio de cultivo utilizado para crecer la cepa portadora de un Cromosoma Artificial Bacteriano (BAC, de sus siglas en inglés) que contenía el *locus* GH de chimpancé, provino de SPECTRUM (Gardena, CA, EUA) y la extracción de ADN BAC se realizó con el Kit BACMAX™ DNA Purification procedente de EPICENTRE (Madison, WI, EUA).

Los reactivos utilizados para la Reacción en Cadena de la Polimerasa (PCR, de sus siglas en inglés) se adquirieron de las siguientes casas comerciales: la enzima DNA polimerasa *Taq* y el $MgCl_2$ de Promega Corporation (Madison, WI, EUA), los dNTPs de New England Biolabs, Inc. (Beverly, MA, EUA) y los iniciadores de Invitrogen (La Jolla CA, EUA) y Sigma Chemical Co. Las amplificaciones por PCR largo se realizaron con el Kit MasterAmp™ Extra-Long PCR de EPICENTRE. La purificación de los productos amplificados de interés, separados por electroforesis en geles de agarosa, fue realizada con el Kit Wizard® SV Gel and PCR Clean-Up System de Promega.

El vector de clonación pTOPO-XL-3.5 fue de Invitrogen y la kanamicina utilizada en los medios selectivos provino de SIGMA. Las enzimas de restricción que se utilizaron en las caracterizaciones enzimáticas proceden de New England Biolabs (Beverly, MA, EUA) y la enzima RNase One™, empleada en el tratamiento de las preparaciones de ADN plasmídico, se obtuvo de Promega corporation. La secuenciación se llevó a cabo con los estuches comerciales de LiCor (LiCor, Lincoln, NE) y ABI Prisma (Perkin Elmer, Cetus, Norwalk, CT).

4.1.2. Material Biológico.

En este trabajo se utilizó ADN procedente de una cepa bacteriana portadora de un BAC, cuyo inserto incluye al *locus* GH del chimpancé, obtenida por tamizaje con una sonda radioactiva del gen hGH de bancos genómicos en BACs del laboratorio del Dr. Pieter de Jong del Children Oakland Hospital Research Institute, en Oakland California, EUA. En la adquisición de este material biológico se contó con la asistencia técnica de Irám Rodríguez.

4.2. Equipo.

Se utilizaron los siguientes equipos: un horno de microondas GoldStar modelo MA-875M, microcentrífuga Eppendorf modelo 5415 (Brinkmann Instruments, Inc. NY, EUA), centrífuga Labnet modelo Spectrafuge 24D (Edison, NJ, EUA) y ultracentrífuga Beckman modelo J2-M1(Beckman Instruments, Inc. CA, EUA), así como un agitador de tubos Barnstead Termolyne modelo Maxi Mix (Bubuque, IOWA, EUA), una placa de agitación magnética con control de temperatura marca Corning modelo PC320 (N.Y. EUA), una incubadora Shell Lab modelo 1535 (Sheldon Manufacturing, EUA), un termomixer Eppendorf (Westbury, NY, EUA) y un baño de agua giratorio modelo G76 (New Brunswick Scientific Co. inc. Edison, N.J. EUA).

Dependiendo de sus requerimientos los reactivos se almacenaron en un congelador marca TorRey de -20°C (Mty, Mex) o en un ultracongelador de -70°C de So Low Enviromental (Cincinnati, Ohio, EUA). La balanza analítica fue de la marca OHAUS modelo AP110S y la granataria de la marca Sartorius modelo 1206MP (Göttingen, Alemania). La electroforesis en geles de agarosa se realizó en cámaras Thermo EasyCast™ B1A y B2 con una fuente de poder BRL 500 (GIBCO-BRL). Para la documentación de la imagen de los geles se utilizó un fotodocumentador UVP Mini Darkroom de Bioimaging Systems (Upland, CA, EUA). El seguimiento del crecimiento bacteriano y la cuantificación del ADN se realizaron con un espectrofotómetro de la marca Eppendorf modelo BioPhotometer. Las amplificaciones por PCR se realizaron en un termociclador Eppendorf Mastercycler gradient (Amburgo, Alemania). El pulso eléctrico para la transformación de las bacterias se hizo en un electroporador modelo Gene Pulser II y celdas de 0.2 mm de la marca BIO-RAD (Hercules, CA, EUA). La secuenciación se realizó en un secuenciador automático ABI Prisma Perkin

Elmmer. El procesamiento de los datos y los análisis bioinformáticos se realizaron en computadoras Macintosh eMac modelo A1002 y laptop Toshiba Satellite T2050.

4.2.1. Programas computacionales.

Los programas de análisis de Biología Molecular utilizados fueron: Amplify versión 3.0 (University of Wisconsin, WI, EUA), OLIGO versión 6.40 (© 1989-2000 Wojciech Rychlik, Molecular Biology Insights, Inc., Cascade, CO, EUA), Mauve 2.1.0 (Darling *et al.*, 2004), GATAligner (Nix y Eisen, 2005), SinicView (Chun-Chieh *et al.*, 2006), CONSED versión 14.0 (University of Washington, Seattle, EUA) (Gordon *et al.*, 1998), DNAMAN versión 5.2.2 Demo (©1994-2001 Lynnon Biosoft Copyright) y MEGA 4.0 (Kumar *et al.*, 2004). El paquete de programas DNASTar Lasergene versión 7.0.0 (Copyright ©1989-2006 DNASTAR Inc., Madison, WI, EUA).

Para la construcción de los vectores y el análisis de patrones de restricción, fue empleado el programa Vector NTI Suite 9 (© 1994-2000 Informax, The University of North Carolina Shared BioInformatics Resource). Para el análisis densitométrico de los productos de PCR se empleó el programa computacional GEL PRO 3.2 (Media Cybernetics, Silver Spring, MD, USA). Para la cuantificación del ADN y la determinación de los tamaños se utilizó el software LabWorks versión 4.5.00.0 (©1996-2003, Media Cybernetics, Inc.).

Los programas bioinformáticos utilizados vía Internet fueron el BLAST disponible en <http://www.ncbi.nlm.nih.gov/BLAST/>, NEBcutter 2.0 de New England Biolabs a disposición en <http://www.neb.com>, ClustalW accesible en <http://www.ebi.ac.uk/Tools/clustalw/>, el programa de predicción de genes GENSCAN disponible en <http://genes.mit.edu/GENSCAN.html> (Standford University, USA), el conjunto de bases de datos y programas para análisis comparativo de secuencias genómicas VISTA disponibles en <http://genome.lbl.gov/vista/index.shtml> y ECR Browser Genome Alignment a disposición en <http://ecrbrowser.dcode.org/genomeAlignment.php>. El programa Repeatmasker Open-3.0 (Smit, AFA., Hubley, R. y Green, P., 1996-2004) para identificación de elementos repetitivos fue utilizado en <http://www.repeatmasker.org/>, mientras que los programas PipMaker y

MultiPipMaker que permiten graficar el alineamiento entre dos o más secuencias genómicas similares fueron empleados en <http://pipmaker.bx.psu.edu/pipmaker/>.

4.3. Métodos.

4.3.1. Extracción de ADN del BAC.

A partir de una cepa de *E. coli* que contenía un BAC portador del *locus* GH del chimpancé, crecida en un medio Luria-Bertani (LB) (bactotripton a 1%, extracto de levadura al 0.5% y NaCl al 1%) adicionado con cloranfenicol, se realizó la extracción de ADN del BAC con el Kit BACMAX™ DNA Purification siguiendo las instrucciones del kit (estuche comercial de EPICENTRE).

4.3.2. Amplificación de genes y regiones intergénicas.

4.3.2.1. Oligonucleótidos diseñados en este trabajo.

Para diseñar los oligonucleótidos se utilizaron los programas Oligo versión 6.40, Amplify versión 3.0 y Vector NTI Suite 9. Los iniciadores utilizados para la amplificación de los genes GH/CSH de chimpancé se diseñaron sobre regiones altamente conservadas de los genes GH conocidos de primates (cuadro 1) y se les incorporaron sitios de restricción (*KpnI* en el 5' y *EcoRV* en el 3') que permitieran su posterior subclonación unidireccional en el vector de expresión pcDNA 3.1 hygro (+).

Cuadro 1. Iniciadores utilizados para amplificar por PCR los genes del *locus* GH del chimpancé.

NOMBRE	SECUENCIA	DESCRIPCION
5'GHPRIMK	AtTaAGGGTaCcGTGGACAGCTCACCTAG	Se une a una secuencia específica de la región 5' no traducible de la unidad transcripcional. Introduce en el amplicón un sitio de restricción <i>KpnI</i>
3'GHPRIMEX	CCtCGaGatatCCTAGAAGCCACAGCT	Se une a una secuencia específica situada entre el quinto exón y la región 3' no traducible de la unidad transcripcional. Introduce en el amplicón un sitio de restricción <i>EcoRV</i> y un sitio <i>XhoI</i>

El diseño de estos iniciadores que amplifican toda la unidad transcripcional de los genes GH/CSH del chimpancé (figura 8) se concibió estratégicamente para que pudiese ser utilizado en cualquier especie de primate y permitir la subsecuente determinación del potencial codificante y funcional de estos genes.

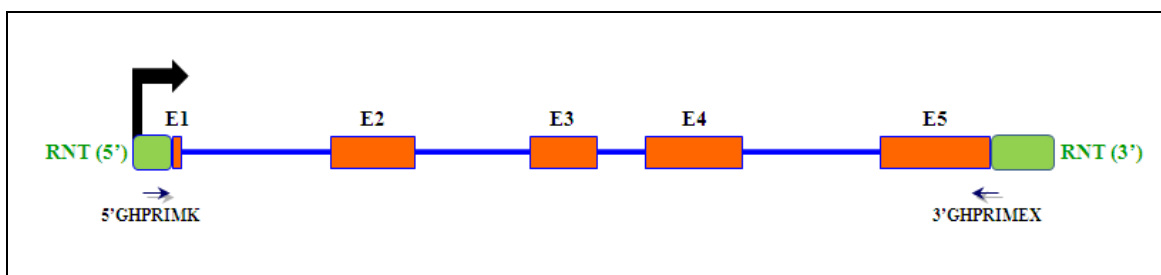


Figura 8. Iniciadores utilizados para amplificar por PCR los genes del locus GH del chimpancé. Se representa la unidad transcripcional de los genes GH/CSH, esquematizándose la estructura exón-intrón típica de estos genes (en color naranja se señalan los cinco exones que conforman estas unidades, separadas por líneas azules que indican los intrones), en verde se representan las regiones no traducibles. Con flechas pequeñas se indican los iniciadores: 5'GHPRIMK y 3'GHPRIMEX.

Para la amplificación del resto del locus de chimpancé se diseñaron series de iniciadores consensos (cuadro 2), ubicados en regiones génicas y regiones intergénicas (figura 9). El diseño de los doce primeros oligonucleótidos mostrados en el cuadro 2, fue basado en la secuencia publicada del *locus* hGH (Chen *et al.*, 1989) y en un ensamblaje preliminar que se realizó al inicio de este trabajo con las secuencias liberadas por el consorcio de secuenciación del genoma del chimpancé en diciembre de 2003 y las secuencias de los genes del chimpancé reportadas por nuestro laboratorio (Revol *et al.*, 2004).

Cuadro 2. Iniciadores utilizados para generar por PCR fragmentos con traslapes que permiten abarcar todo el locus GH del chimpancé.

	NOMBRE	SECUENCIA		NOMBRE	SECUENCIA
01	5'GH/CS-Con	CATCTCCCTGCTGCTCAT	12	3'TECAM-Con	CCATAGCCACTTTTGATGC
02	3'IGC-Con	GGGGAATTGAACACAAGGA	13	FDLCHIM	GCTCATGCCAGTAATCTCAGTMC
03	5'PlacenCon	GGAGCTCAGGGTTGTTTTTC	14	RDLCHIM	GGTTCAGACACCATTGAGTC
04	5'IGC-Con	TCTCCCTGGAGAATAATCCC	15	FL1PA3	TGCTGCACCCATTAACCTTGTC
05	3'CHElemPCon	GTAAGACGGGGTTTCACTATG	16	IGLEcoF	TTTGACAGATACAGGATATCTACAGCCCTGATG
06	5'CHElemPCon	AGCAGGGGAAAGTGGTGA	17	IGLKpnR	CCCCACAGTTTGGTACCCTAAGATTTAGGACTAC
07	3'PlacenCon	GATGAGCAGCAGGGAGATG	18	ENHF	CTGAGATTCTGATATAATTAGCCTGGAATGTGG
08	3'IGL-Con1	CTTTTCTGCTCCTTCCC	19	ENHR	CACACCAAGTTGATGAGCTGAAATGTAGACTAAC
09	5'IGL-Con1	GAGAGCAGAACTGACCCCTTC	20	ELEMPF	GTTGCCAACACCACAGCCAACCAC
10	3'IGL-Con2	GGTAACCTTTGCTGCCCTG	21	ELEMPR	CTGTGGCCTAGGATGTGAATGCTGG
11	5'IGL-Con2	GCTTTCCTCCTCTTTGCTCT			

Los restantes iniciadores se diseñaron en base a una actualización de dicho ensamblaje con las secuencias que fueron obteniéndose en este trabajo y con los datos de secuencias procedentes de la actualización liberada en octubre de 2006 por el consorcio de secuenciación del genoma del chimpancé.

En la parte superior de la figura 9 se representan los amplicones que incluyen un gen relacionado a GH y regiones intergénicas contiguas. En la parte inferior se representan aquellos amplicones obtenidos con un iniciador que se une en una región interna de los genes GH/CSH y su pareja que es complementaria a una región intergénica, a excepción del fragmento amplificado con los iniciadores 01 y 12, en el cual el segundo iniciador de este par es complementario a una región interna del gen TCAM.

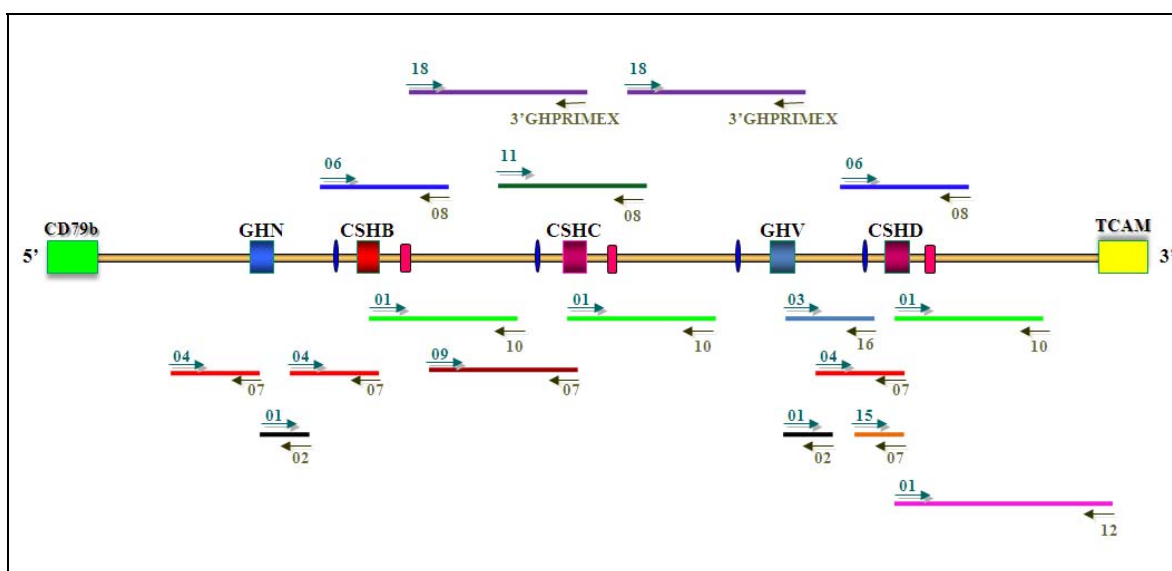


Figura 9. Iniciadores utilizados para amplificar el locus GH del chimpancé. Los productos amplificados se representan con líneas horizontales que se traslapan abarcando todo el locus GH del chimpancé. Con flechas horizontales ubicadas sobre los extremos de los productos amplificados representados se indican los iniciadores utilizados.

4.3.2.2. Condiciones de amplificación de los genes del chimpancé.

Los genes GH/CSH del locus del chimpancé fueron amplificados por PCR utilizando el par de iniciadores específicos mostrados en el cuadro 1. La reacción de PCR se llevó a cabo utilizando DNA polimerasa Taq Gold (EPICENTRE). Se realizaron múltiples ensayos a fin de optimizar las condiciones de PCR para la amplificación de los genes GH/CSH de esta

especie. El resultado de esta estandarización se muestra en el cuadro 3. La mezcla de reacción fue ajustada con agua ultrapura a un volumen final de 50 μ l.

Cuadro 3. Condiciones de amplificación de los genes *GH/CSH* del chimpancé.

Componente de la PCR	Concentración final
Amortiguador 10 X	1 X
MgCl ₂ 25 mM	1.5 mM
dNTPs 10 mM	200 μ M
Oligo sentido 5 μ M	0.2 μ M
Oligo anti-sentido 5 μ M	0.2 μ M
Taq polimerasa 5 U/ μ l	1 U
ADN BAC	10 ng
Agua ultrapura	cbp 50 μ l

Después de un paso inicial de desnaturalización a 94 °C por 5 minutos, las amplificaciones se llevaron a cabo en 30 ciclos, consistentes cada uno de un paso de desnaturalización a 94 °C por 1 minuto, uno de alineamiento a 64 °C por 1 minuto y un tercero de extensión a 72 °C por un minuto. Terminados los 30 ciclos, se programó un paso final de extensión a 72 °C durante 5 minutos. El producto amplificado se visualizó en un gel de agarosa al 1%, seguido de tinción con bromuro de etidio y exposición a luz ultravioleta. Posteriormente fue caracterizado con diferentes enzimas de restricción para corroborar su identidad.

4.3.2.3. Condiciones de amplificación de las regiones intergénicas del chimpancé.

La estandarización de la PCR larga se realizó utilizando cada una de las nueve Pre-mezclas 2X (contiene amortiguador Extra-Long PCR 2X, 400 mM de cada dNTP, concentraciones de Mg²⁺ de 3 a 7 mM y betaína) que suministra el fabricante y variando las concentraciones de ADN genómico o BAC en la mezcla de reacción de la PCR, temperatura de alineamiento de los iniciadores y los tiempos de extensión, logrando así establecer como óptimas las condiciones de reacción aquí descritas. El cuadro 4 describe la composición de la mezcla de reacción con un volumen final de 50 μ l y el programa de amplificación se detalla en el cuadro 5. Los productos amplificados, junto con un marcador de peso molecular, se corrieron en un gel de agarosa al 0.8% para estimar su tamaño y

posteriormente se caracterizaron con diferentes enzimas de restricción para identificar su origen.

Cuadro 4. Composición de las mezclas de reacción de la PCR larga.

Componente de la PCR	Concentración final
Oligo sentido 5 μ M	0.2 μ M
Oligo anti-sentido 5 μ M	0.2 μ M
Taq polimerasa 2.5 U/ μ l	2.5 U
ADN de BAC	10 ng
Agua ultrapura	cbp 25 μ l
Pre-mezcla 2X # 4	1X

Cuadro 5. Programa de amplificación de la PCR larga.

1. Desnaturalización inicial	94 °C por 5 minutos
2. 12 ciclos	
Desnaturalización	94 °C por 30 segundos
Alineamiento	55 - 57 °C por 40 segundos
Extensión	72 °C por 7-14 minutos
3. 23 ciclos	
Desnaturalización	94 °C por 30 segundos
Alineamiento	55 - 57 °C por 40 segundos
Extensión	72 °C por 7-14 minutos con incrementos de 15 segundos cada ciclo
4. Extensión prolongada	72 °C por 20 minutos

4.3.3. Clonación de los genes y regiones intergénicas.

Los productos amplificados del *locus* del chimpancé se clonaron en el vector comercial pTOPO-XL-3.5 de Invitrogen, de acuerdo a las indicaciones del fabricante TOPO® XL PCR Cloning Kit.

4.3.3.1. Preparación de células calcio competentes.

Las bacterias DH5 α calcio competentes fueron preparadas de la siguiente forma: una colonia de bacterias fue inoculada en 4 ml de medio LB. Las bacterias fueron crecidas durante toda la noche a 37 °C en agitación constante. Al día siguiente, fueron inoculados 0.5 ml del cultivo bacteriano en 50 ml de medio LB fresco y puestas a crecer a 37 °C, con agitación constante hasta alcanzar una DO (densidad óptica) de 0.375 a 590 nm. Una vez

alcanzada la densidad óptica deseada, las bacterias fueron centrifugadas a 3,000 rpm durante 7 minutos. Luego de eliminado el sobrenadante, la pastilla bacteriana fue resuspendida en 10 ml de CaCl_2 frío. Posteriormente fueron sometidas a una nueva centrifugación a 2,500 rpm durante 5 minutos. Luego de remover el sobrenadante fueron resuspendidas en 10 ml de CaCl_2 frío y mantenidas en hielo durante 30 minutos. Alícuotas de 200 μl fueron preparadas al término del procedimiento y preservadas a -70°C .

4.3.3.2. Transformación de células calcio competentes.

Para la transformación fueron empleadas 200 μl de las bacterias tratadas con CaCl_2 , las cuales fueron incubadas durante 30 min. a una temperatura de 4°C con los productos de la reacción de ligación obtenida para los genes GH/CSH de chimpancé. Posteriormente se realizó un choque térmico durante 1 min. a una temperatura de 42°C , seguido de una incubación en hielo. Para recuperar las bacterias del choque térmico se agregaron 800 μl de medio LB y se incubaron durante 1 hora a una temperatura de 37°C . Finalmente, las bacterias fueron centrifugadas a 3,000 rpm durante 5 minutos, se descartó el sobrenadante y la pastilla constituida de las bacterias se resuspendió en 200 μl de medio LB, para finalmente ser sembradas en placas de Petri conteniendo medio LB-Agar con antibiótico kanamicina como medio de selección e incubadas a 37°C durante 16 horas.

4.3.3.3. Tamizaje de las clonas recombinantes portadoras de los genes GH/CSH.

Las células calcio competentes transformadas fueron tamizadas por PCR en colonia con los iniciadores apropiados para confirmar la presencia de los genes GH/CSH de chimpancé. Los productos amplificados fueron analizados mediante electroforesis en geles de agarosa al 1 %, utilizando como marcador de peso molecular al bacteriófago λ digerido con la enzima de restricción *Pst*I. Posteriormente los amplicones fueron caracterizados con enzimas de restricción (*Pst*I, *Bst*EII y *Avr*II) y los patrones de digestión fueron comparados con los esperados por predicción en base a sus secuencias. Las clonas que resultaron positivas fueron crecidas a 37°C en tubos con LB suplementado con kanamicina como medio selectivo y en agitación constante durante 18 horas. Finalmente, se procedió a la extracción del ADN plasmídico de las clonas recombinantes identificadas.

4.3.3.4. Transformación de células electrocompetentes.

La electroporación de las bacterias *E. coli* TOP 10 “One Shot” de Invitrogen se hizo con 2 μ l de los productos de la reacción de ligación obtenida para los fragmentos del *locus* GH del chimpancé amplificados por PCR larga. Las condiciones de electroporación para transformar las bacterias fueron: 200 Ω , 25 μ F y 2.5 kV. Después del pulso se transfirieron las bacterias a 450 μ l de medio SOC (incluido en el TOPO® XL PCR Cloning Kit) y se incubaron a 37 °C con agitación (800 rpm) durante una hora. Posteriormente 100 μ l de la suspensión bacteriana se sembraron en placa en LB suplementada con kanamicina (50 mg/ml), misma que se incubó a 37 °C durante 16 horas. Pasado este tiempo, se pasaron las colonias a tubos con LB igualmente suplementado con kanamicina y se pusieron a crecer a 37 °C en agitación constante durante 18 horas. Finalmente, se procedió a la extracción del ADN plasmídico.

4.3.3.5. Extracción del ADN plasmídico.

Se centrifugaron 3 ml del cultivo bacteriano durante 5 min a 5,000 rpm y se removió el medio por aspiración, manteniendo el tubo en hielo. El ADN plasmídico se extrajo por el método de la lisis alcalina (Sambrook *et al.*, 2001). Las células se resuspendieron en 100 μ l de solución I (glucosa a 50 mM, Tris-HCl a 25 mM y EDTA a 10 mM y pH 8) con el vortex. Se agregaron 200 μ l de solución II (preparada en fresco mezclando 0.2 ml de NaOH a 10 N y 1.0 ml de SDS al 10% en 8.8 ml de agua), mezclando inmediatamente por inversión e incubando en hielo por 5 min. Pasado este tiempo, se agregaron 150 μ l de solución III (acetato de amonio a 7.5 M), se mezcló de inmediato por inversión y se incubó por 5 min en hielo. El tubo se centrifugó a 14,000 rpm por 5 minutos y el sobrenadante se recuperó en un tubo nuevo donde se le hizo una extracción fenol/Sevag (½ volumen de fenol y ½ de Sevag). Se centrifugó a 14,000 rpm por 2 min. y el sobrenadante se transfirió a un tubo nuevo donde el ADN plasmídico se precipitó agregando dos volúmenes de etanol al 100% frío y mezclando por inversión completamente. Después de centrifugar a 14,000 rpm por 5 min, la pastilla se lavó dos veces con 1 ml de etanol al 70% previamente enfriado y se dejó secar a temperatura ambiente durante 30 minutos. Finalmente, se resuspendió en 50 μ l de TE (pH 8.0) conteniendo RNasa a 20 μ g/ml.

4.3.3.6. Selección y análisis de las clonas recombinantes portadoras de los fragmentos del *locus* GH del chimpancé amplificados por PCR larga.

Los plásmidos fueron analizados mediante electroforesis en geles de agarosa al 0.8 %, utilizando como marcador de peso molecular al bacteriófago λ digerido con la enzima de restricción *Pst*I. Los plásmidos recombinantes que mostraron un retraso en su patrón de corrimiento en correspondencia con el tamaño del inserto, fueron seleccionados como candidatos y posteriormente para caracterizarlos se sometieron a caracterizaciones con enzimas de restricción (*Pst*I, *Bst*EII y *Avr*II). Los amplicones que originaron las clonas fueron digeridos en idénticas condiciones y corridos en el mismo gel. Los patrones de corrimientos fueron comparados con los esperados por los análisis *in silico* por predicción en base a sus secuencias.

4.3.4. Análisis de las secuencias.

4.3.4.1. Secuencias utilizadas en este estudio.

Con la secuencia NW_118509 liberada por el consorcio de secuenciación del genoma del chimpancé en diciembre de 2003 y las secuencias (AY146628-AF374232) de los genes del chimpancé reportadas por nuestro laboratorio (Revol *et al.*, 2004) se realizó al inicio de este trabajo un ensamblaje preliminar del *locus* GH del chimpancé utilizando el programa Genestudio y tomando como referencia la secuencia J03071 correspondiente al *locus* hGH.

Cuadro 6. Secuencias utilizadas en este estudio.

Especie	No. de Acceso
Humano (<i>Homo sapiens</i>)	
<i>Locus</i> GH/CSH	J03071
Chimpancé (<i>Pan troglodytes</i>)	
GH-N	AF374232 (AAL72285)
GH-V	AF374233 (AAL72285)
CSH-A	AY146625 (AAN84505)
CSH-B	AY146626 (AAN84506)
CSH-C	AY146627 (AAN84507)
CSH-D	AY146628 (AAN84508)
Cerdo (<i>Sus scrofa</i>)	
GH	DQ073397 (AAY59032)
Proyecto genómico del chimpancé	
Secuencias de Diciembre de 2003	NW_118509
Actualización de Octubre de 2006	NW_001227001
	AACZ02172434 - AACZ02172445

Posteriormente se realizó una actualización de dicho ensamblaje con las secuencias que fueron obteniéndose en este trabajo y con los datos de secuencias (NW_001227001, AACZ02172434 - AACZ02172445) procedentes de la actualización liberada en octubre de 2006 por el consorcio de secuenciación del genoma del chimpancé.

4.3.4.2. Identificación de las secuencias generadas.

Las secuencias de genes y regiones intergénicas generadas en este estudio se compararon en el BLAST (Altschul, *et al.*, 1990) con las secuencias disponibles en las bases de datos del GenBank, con el fin de determinar su identidad. Adicionalmente fueron analizadas con el programa vecscreen para verificar que no hubiese contaminación con secuencias del vector. Finalmente se compararon con los ensamblajes genómicos disponibles para el chimpancé y el humano, con el propósito de determinar su probable ubicación y facilitar el reensamblaje y finalización del mismo en el chimpancé, integrando toda esta información.

4.3.4.3. Ensamblaje de los *loci* GH.

Las secuencias de los genes y las regiones intergénicas generadas en este estudio fueron ensamblándose con el programa CAP3 (Huang, X. y Madan, A., 1999) en la medida en que se fueron obteniendo para determinar su organización. Se utilizaron adicionalmente las secuencias reportadas por el consorcio de secuenciación del genoma del chimpancé y como referencia el *locus* GH del humano.

4.3.4.4. Análisis de las secuencias nucleotídicas de los genes.

Las secuencias nucleotídicas de los genes obtenidas en este trabajo se alinearon con las reportadas para los genes GH/CSH del humano (cuadro 6) con el programa MEGALIGN del paquete de programas DNASTar Lasergene versión 7.0.0 por el método ClustalW para determinar la similitud entre ellas. Finalmente se determinaron las relaciones filogenéticas entre dichas secuencias por el método de Neighbor-Joining (NJ) (Saitou *et al.*, 1987) con un valor de Bootstrap de 1000 réplicas mediante el programa MEGA 3.0 (Kumar *et al.*, 2004) utilizando como grupo fuera la secuencia del cerdo.

4.3.4.5. Análisis de las proteínas.

Las secuencias aminoacídicas se dedujeron de las nucleotídicas mediante el programa Genscan (Burge y Karlin, 1997), (Burge, 1998) y (Burge y Karlin, 1998) y con las primeras se hizo el alineamiento en el programa MEGALIGN del paquete de programas DNASTar Lasergene versión 7.0.0 por el método ClustalW para determinar la similitud entre ellas e identificar los aminoácidos clave en la unión al receptor (Liu *et al.*, 2001).

4.3.4.6. Análisis del promotor proximal.

El análisis se centró en identificar los sitios de unión a factores transcripcionales descritos como responsables de la transcripción basal del gen hGH-N y se enfocó en la región que va del nucleótido -134 al -1. Dentro de estos destacan Pit-1 proximal y distal, la caja TATA y los elementos SP1 e InrE. Las secuencias de cada uno de los promotores proximales de los genes GH/CSH de chimpancé se alinearon con su contraparte en el humano utilizando el programa MEGALIGN del paquete de programas DNASTar Lasergene versión 7.0.0 por el método ClustalW.

4.3.4.7. Análisis de los elementos reguladores del control transcripcional.

El análisis consistió en identificar con el programa Bl2Seq (Tatusova y Madden, 1999) los dos elementos claves de la especificidad tisular: el inhibidor o elemento P que reprime la expresión de los genes placentarios en la hipófisis (ubicado 2 kpb río arriba de dichos genes) y el potenciador o “enhancer” que incrementa la expresión de los genes CSHs en la placenta (ubicado a 2.3 kpb pero río debajo de los genes CSHs). Las secuencias de cada uno de los elementos reguladores del control transcripcional se alinearon con su contraparte en el humano utilizando el programa MEGALIGN del paquete de programas DNASTar Lasergene versión 7.0.0 por el método ClustalW.

4.3.4.8. Identificación de los elementos repetitivos.

La identificación de los elementos repetitivos se realizó en las secuencias de las regiones intergénicas del chimpancé y el humano con el programa RepeatMasker (Smit, AFA., Hubley, R. y Green, P., 1996-2004).

CAPÍTULO V: RESULTADOS

5.1. Iniciando el análisis de los datos sobre el *locus* GH del genoma del chimpancé.

Este trabajo persiguió ensamblar el *locus* GH del chimpancé (figura 10), partiendo del archivo NW_118509 liberado por el Consorcio de Secuenciación del Genoma del Chimpancé (CSGCh) en diciembre de 2003 y de las secuencias (AY146628-AF374232) de los genes de dicho *locus* clonados en nuestro laboratorio (Revol *et al.*, 2004). Inició apoyándose en el programa bioinformático Genestudio y en la secuencia del *locus* hGH, como referencia (número de acceso J03071).

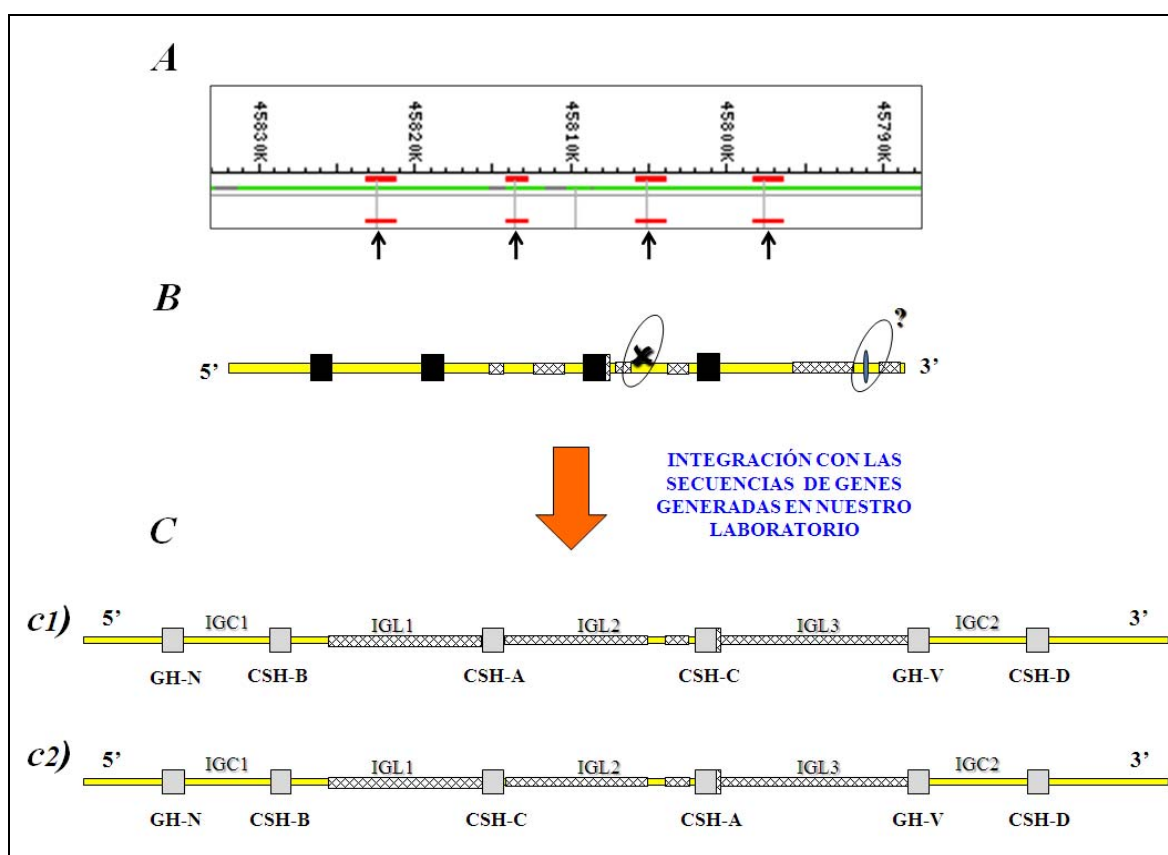


Figura 10. Ensamblaje preliminar de la región correspondiente al *locus* GH del chimpancé. En A se muestra la región génica del *locus* GH del chimpancé deducida a partir de la información liberada en diciembre de 2003 por el CSGCh (con flechas se representan las regiones identificadas como genes, por su similitud al gen hGH-N en la búsqueda bioinformática con BLAST). En B se representa un esquema del estado de la secuenciación en dicha región y en C dos modelos propuestos (c1 y c2) para el *locus* GH del chimpancé, en base a los datos de secuencias del consorcio y los correspondientes reportados por nuestro laboratorio. La principal diferencia entre estos dos modelos radica en el orden de los genes CSH-A y CSH-C. Con cuadros negros se representan las secuencias del CSGCh similares a GH, con cuadros grises los genes GH reportados por nuestro laboratorio, con un óvalo y una X se señala las secuencias no relacionadas con el *locus* GH, con un óvalo y un ? se indica las secuencias de dudosa ubicación y con un rayado de diagonales se representan los huecos en la secuencia que se tenían en ese entonces.

Un primer paso fue determinar correspondencias entre ambas clases de secuencias, encontrándose que en las primeras había cuatro regiones que correspondían con cuatro de los seis genes en las segundas, exhibiendo similitudes que oscila entre 96 y 99 (ver el cuadro 7). El valor de 96% para el gen CSH-B fue un resultado inesperado, pero pudo explicarse por los errores de ensamblajes y huecos presentes en la secuencia liberada por el CSGCh.

Cuadro 7.- Comparación entre genes GH/CSH de nuestro laboratorio y del CSGCh.

GENES DE CSGCh RELACIONADOS A GH	GENES DESCUBIERTOS EN NUESTRO LABORATORIO	SIMILITUD
Similar a GH-N	GH-N	99 %
Similar a CSH-B	CSH-B	96 %
Similar a GH-V	GH-V	99 %
Similar a CSH-D	CSH-D	98 %
-	CSH-A	-
-	CSH-C	-

Dos genes (CSH-A y CSH-C) de los reportados por nuestro equipo de trabajo no estaban representados en la secuencia del correspondiente “contig” (Secuencias contiguas de ADN originadas por el ensamblaje de fragmentos secuenciados de un genoma) en el archivo del GenBank del CSGCh, por lo que dedujimos que forman parte de los huecos que se hallan en dicha secuencia. Nuestra inspección del *locus* GH del consorcio, experiencia en el mismo y la inclusión de regiones intergénicas, nos permitió deducir la primera estructura del *locus* presentada en la figura 10.

5.2. Actualización del ensamblaje del *locus* GH propuesto.

En octubre de 2006 el CSGCh libera una actualización de las secuencias genómicas previamente reportadas. Dicha actualización se parece a la estructura obtenida en el ensamblaje preliminar realizado al inicio de este trabajo. En base a dicha actualización (NW_001227001, AACZ02172434 - AACZ02172445) se procedió en nuestro laboratorio a hacer una revisión (figura 11) del ensamblaje preliminar, considerando nuevamente tanto secuencias que habían sido obtenidas en el presente trabajo, como los datos nuevos de secuencias actualizadas por el CSGCh. Adicionalmente se analizaron zonas de conflictos

en el ensamblaje y se verificaron alternativas para ensamblar estas regiones utilizando las secuencias trazas (NCBI) y como referencia el *locus* humano.

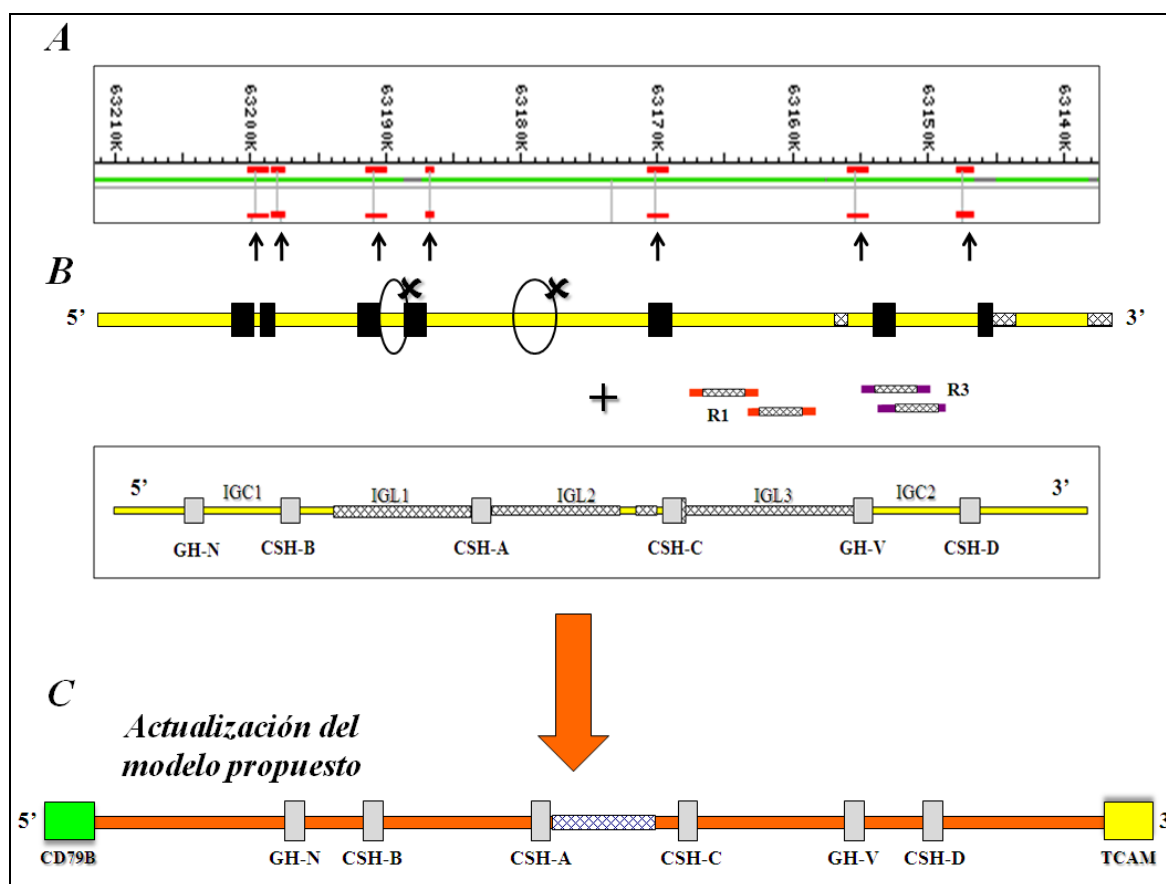


Figura 11. Actualización de las secuencias en la región génica del *locus* GH del chimpancé. En **A** se muestra la región génica del *locus* GH del chimpancé interpretada a partir de la actualización liberada en octubre de 2006 por el CSGCh (con flechas negras se representan las regiones identificadas por su similitud en BLAST con la secuencia nucleotídica del gen hGH-N). En **B** se representa un esquema del estado de la secuenciación del consorcio en dicha región, el modelo más probable propuesto para el *locus* GH de chimpancé y las primeras secuencias obtenidas en este trabajo (R1 y R3). En **C** se representa la actualización del modelo propuesto para el *locus* GH después de integrar los datos de secuencias. Con cuadros negros se representan las secuencias similares a GH, con cuadros grises los genes GH reportados por nuestro laboratorio, con un óvalo y una **x** se señala las secuencias no relacionadas con el *locus* GH y con un rayado de diagonales cruzadas se representan los huecos en la secuencia.

A pesar de actualizar el ensamblaje, aun quedaban huecos en regiones intergénicas y dudas en el orden de los genes CSH-A y CSH-C, pues los análisis no encuadraban bien. Además, por más que se intentó, no fue posible amplificar el supuesto gen CSH-A y coincidentemente con los datos de secuencias del consorcio, tampoco encontramos secuencia alguna que pudiese corresponder con dicho gen. Teniendo en cuenta esta situación se hizo necesario considerar la posibilidad de que el *locus* pudiese estar formado

por cinco genes en vez de seis, para lo que se propusieron dos modelos de ensamblaje. Particularmente valiosas resultaron las diferencias entre ambos modelos (el de cinco y el de seis genes) a nivel de las intergénicas largas, el arreglo de cinco genes tendría dos intergénicas largas, mientras que el de seis tendría tres. Basados en los ensamblajes propuestos para los dos arreglos de genes, se predijeron *in silico* mapas de restricción con el objetivo de recurrir a ello para verificar a cuál arreglo corresponderían los resultados experimentales.

Debido a que el ensamblado correspondiente al modelo de seis genes contenía huecos de secuencias y por simplicidad, la hipótesis a probar primero fue el modelo de cinco genes. En caso de que el patrón de cortes con enzimas de restricción sobre las clonas de los genes y de las regiones intergénicas de interés generase fragmentos con pesos moleculares cuya suma arrojara valores mayores a los esperados, se optaría luego por el modelo de seis genes.

En base a las generalidades encontradas en el *locus* GH en MVM y en el humano, se utilizaron iniciadores en zonas consensos específicos para intergénicas largas y otros para intergénicas cortas. El tamaño aproximado de las secuencias faltantes en el caso del arreglo de seis genes sería posible estimarse en base a los pesos moleculares de los productos amplificados “extras”. El dilema se resolvió pronto, pues los patrones de cortes obtenidos, tanto para los productos amplificados como para las clonas derivadas, siempre fueron consistentes con las predicciones realizadas para el arreglo de cinco genes. En base al modelo comprobado inició la caracterización del *locus*.

5.3. Aclarando dudas sobre los genes.

El ADN extraído a partir de una cepa de *E. coli* que contenía un BAC portador del *locus* GH del chimpancé, como se deduce de su perfil electroforético (figura 12A), resultó de muy buena calidad. A partir de este se amplificaron los genes GH/CSH del chimpancé y los amplicones generados fueron verificados observándose una banda única de 1533 pb, que coincide con el peso molecular esperado para dichos productos (figura 12B).

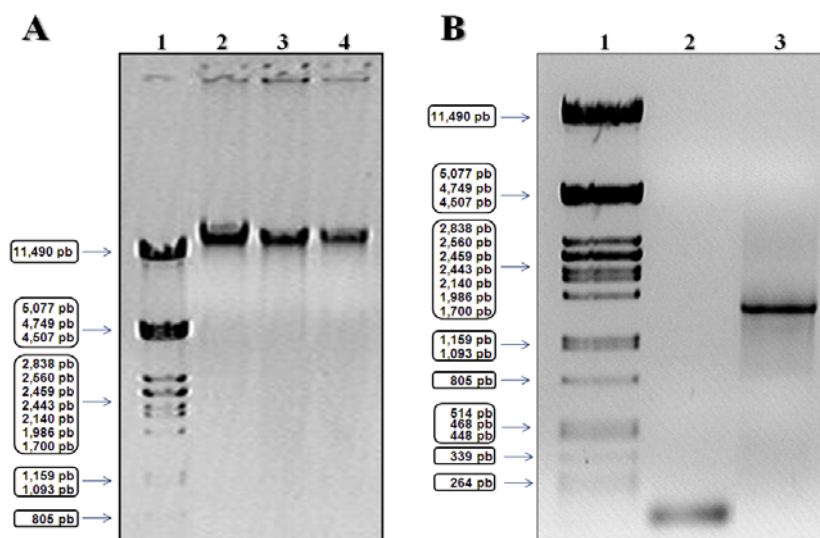


Figura 12. Verificación de la integridad del ADN del BAC y de la amplificación de los genes del locus GH del chimpancé. Electroforesis en gel de agarosa al 1%. En **A**, carril 1: como marcador de peso molecular se utilizó el ADN del bacteriófago λ digerido con la endonucleasa de restricción *Pst*I; carriles 2, 3 y 4: ADN de BAC obtenido de las extracciones realizadas. En **B**, en el carril 1 se utilizó el mismo marcador de peso molecular que el empleado en **A** (como también lo será en lo sucesivo identificándose simplemente como λ), en el 2 el control negativo para la PCR y en el 3 el producto de 1533 pb amplificado con el par de iniciadores consenso 5'GHPRIMK y 3'GHPRIMEX, para genes GH/CSH.

Los productos amplificados se clonaron y las clonas positivas fueron tamizadas mediante PCR en colonia con los iniciadores empleados para amplificar las unidades transcripcionales de los genes GH/CSH (figura 13). Se observó una banda única de 1533 pb que delató a aquellas clonas que contenían vectores portando genes GH/CSH.

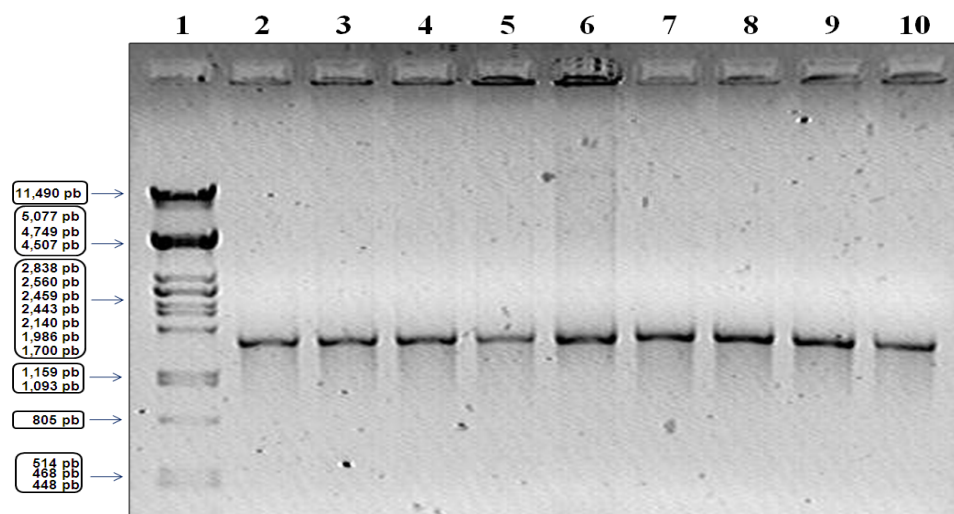


Figura 13. Tamizaje mediante PCR en colonia de vectores recombinantes portadores de GHs. Electroforesis en gel de agarosa al 1%, en el cual fueron identificadas las clonas que contenían los vectores portadores de las unidades transcripcionales de GHs. En el carril 1 el marcador λ y en el resto de los carriles las PCRs de clonas.

Posteriormente el producto amplificado obtenido fue caracterizado mediante cortes con enzimas de restricción y los patrones de corte obtenidos fueron comparados con los esperados por los análisis *in silico* de predicción en base a las secuencias nucleotídicas correspondientes.

La digestión enzimática con *Pst*I permitió distinguir a los genes GH de los CSH. Como se observa en la figura 14, las muestras analizadas en los carriles 3, 6, 8, 16 y 17 exhiben patrones de corte que corresponden con los esperados para los genes GH. Sin embargo, el patrón del carril 6 permite identificar a esta muestra como un gen GH-V, no obstante como las diferencias entre el patrón de corte de GH-N y GH-V son sutiles, se corroboró el resultado con un segundo ensayo de restricción. El resto de las clonas exhibieron cortes compatibles con genes CSHs.

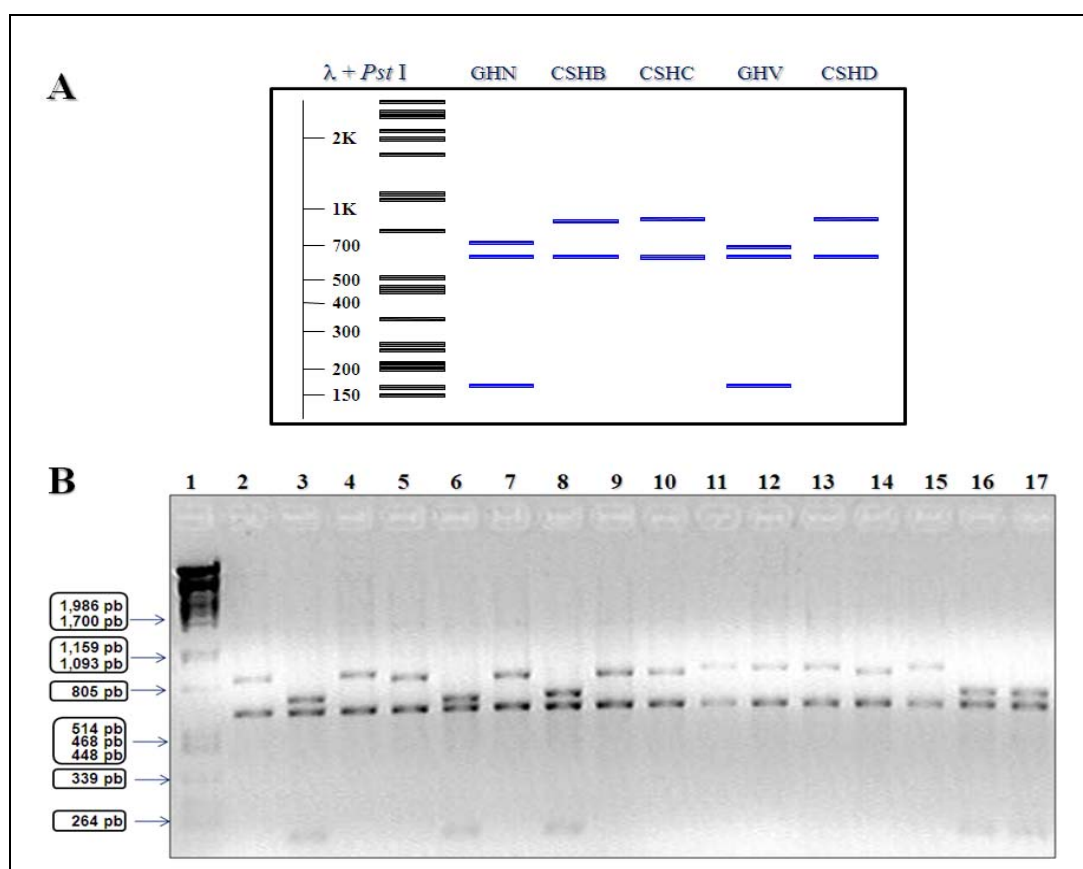


Figura 14. Caracterización de los productos amplificados por digestión con *Pst*I. En **A** se muestran los patrones de digestión esperados en base a las secuencias. En **B** se incluye un gel de agarosa al 2% en el cual se verificaron los productos de digestión. Como marcador de peso molecular se utilizó λ (carril 1). Los carriles del 2 al 17 muestran los patrones obtenidos al digerir con *Pst*I los productos amplificados a partir de las clonas portadoras de genes GH/CSH de chimpancé.

Una segunda caracterización fue realizada mediante cortes con la enzima de restricción *Bam*HI, para distinguir específicamente a los genes GH-V y GH-N (figura 15). Como se observa en 15A, el carril 3 exhibe un patrón de corte que permite identificar al amplicón procedente del gel como GH-V, mientras que el resto de las muestras digeridas con *Bam*HI coinciden con los resultados esperados para los genes GH-N. Los genes identificados fueron secuenciados, para no dejar duda alguna de su identidad.

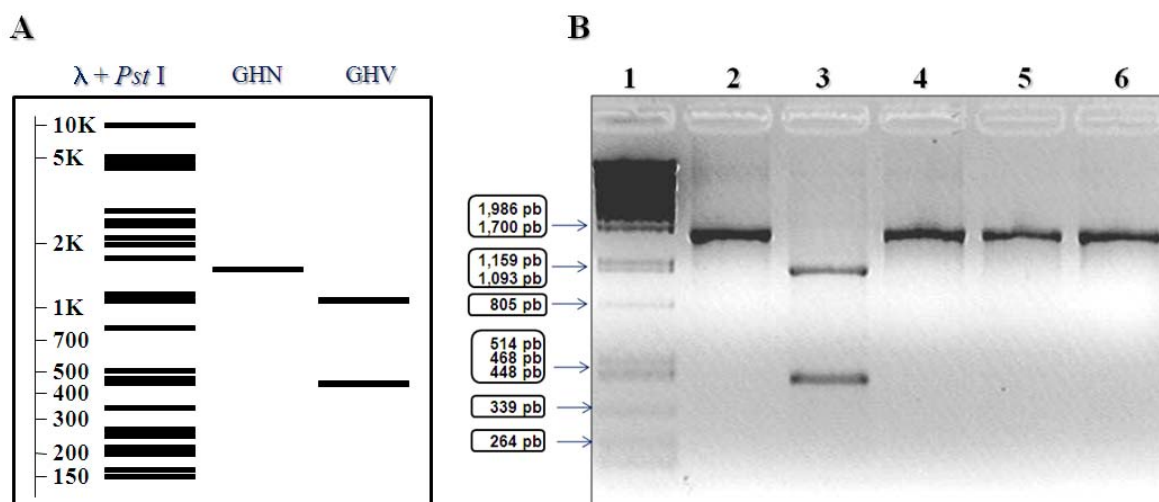


Figura 15. Caracterización de los productos amplificados por digestión con *Bam*HI. A la izquierda se muestra los patrones de digestión esperados para la enzima *Bam*HI. A la derecha se muestra un gel de agarosa al 2% en el cual se verificaron los productos de digestión. En el carril 1 se incluyó el marcador de peso molecular correspondiente al ADN del bacteriófago λ digerido con la enzima *Pst*I. Los carriles del 2 al 6 muestran los patrones obtenidos al digerir con *Bam*HI los productos amplificados a partir de las clonas portadoras de los genes GH de chimpancé.

La diferenciación de los genes CSH no se realizó con enzimas de restricción, debido a que no fue posible obtener patrones de restricción excluyentes que permitieran su identificación sin lugar a dudas. Resultó más práctico aislarlos a partir de fragmentos clonados que incluían el gen con sus regiones intergénicas flanqueantes y que facilitaron la determinación de su identidad como se muestra más adelante en la sección 5.6.

5.4. Las regiones intergénicas como claves en el arreglo de los genes.

Buscando resolver la contradicción de los modelos de cinco y seis genes, se recurrió a la amplificación de las regiones intergénicas a partir de secuencias conservadas en dichas regiones y en los genes. Por medio de la PCR larga, con múltiples combinaciones de iniciadores y utilizando como templado el ADN del BAC portador del *locus* GH del

chimpancé, se obtuvieron productos amplificados con traslapes de forma tal que permitieran cubrir completamente el *locus*. Ejemplos representativos se muestran en la figura 16.

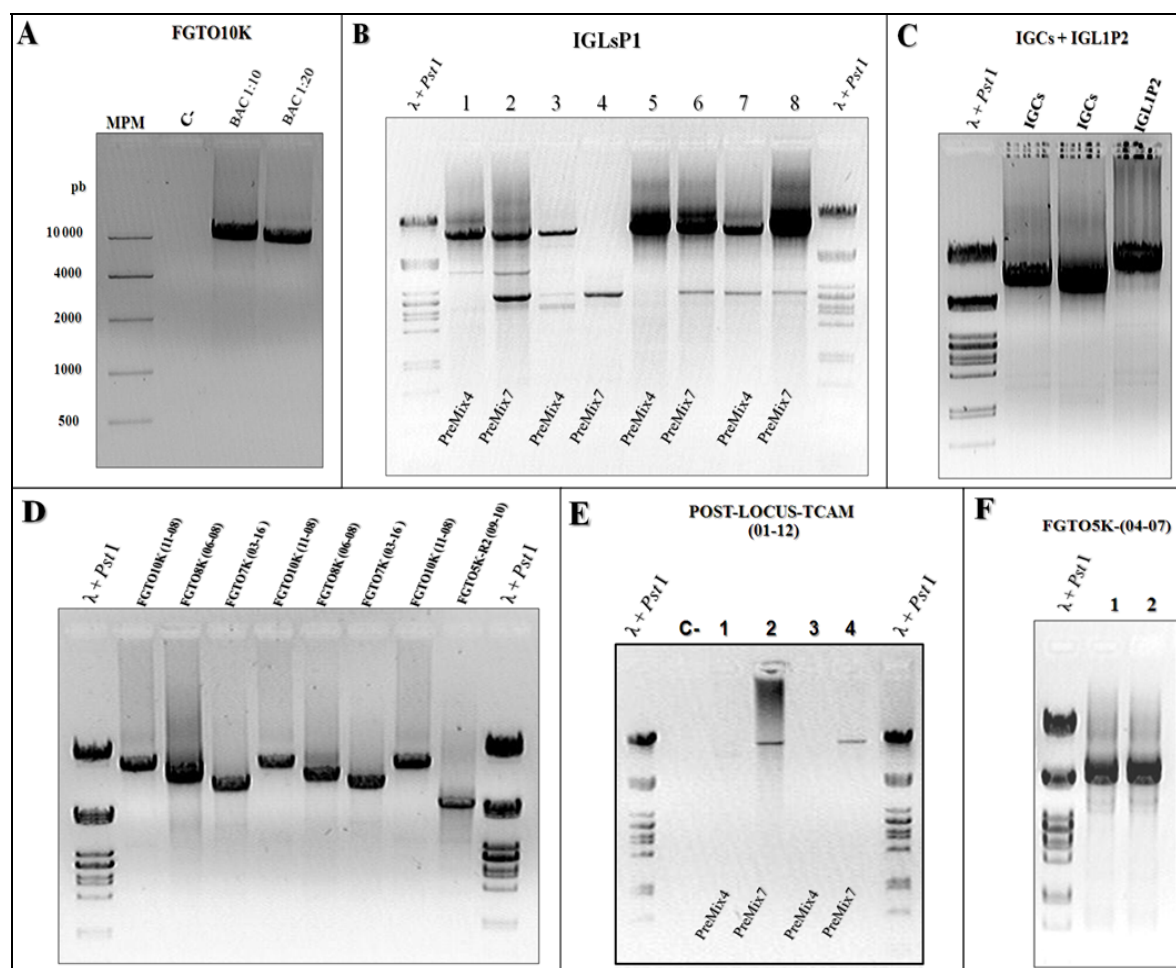


Figura 16. Amplificación de las regiones intergénicas del *locus* GH del chimpancé. Los productos de amplificación se resolvieron en geles de agarosa al 1%. **A.** Productos amplificados identificados como FGTO10K, que abarcan regiones de dos intergénicas largas consecutivas y el gen incluido entre ellas. **B.** Productos de amplificación que abarca los primeros dos tercios de las intergénicas largas y de la región posterior al último gen del *locus*. **C.** Producto amplificado correspondiente a las intergénicas cortas (IGCs) obtenida con los iniciadores GHPRIMF - GHPRIMR y a otra región de la primera intergénica larga (IGLIP2) con los oligonucleótidos 09 – 07. **D.** Productos de amplificaciones correspondientes a diferentes regiones del *locus*. **E.** Producto amplificado comprendido entre la segunda mitad del último gen del *locus* y secuencias más allá de éste (TCAM). **F.** Productos de amplificación que abarcan las regiones anteriores a los genes GH-N, CSH-B y CSH-D incluyendo la primera mitad de la secuencia de estos. Con la excepción del marcador en A (escalera de 500 a 10,000 pb) el marcador fue λ .

Todos los productos amplificados obtenidos fueron sometidos a caracterización con enzimas de restricción y los resultados experimentales comparados con las respectivas predicciones en base a secuencias nucleotídicas. Adicionalmente fueron clonados y

finalmente las clonas portadoras de los mismos fueron caracterizadas de forma independiente con las enzimas *AvrII*, *BstEII* y *PstI*.

5.5. La región crítica para resolver el dilema del arreglo génico.

En la figura 17 se muestran los dos posibles arreglos de genes que se propusieron en este trabajo para el *locus* GH del chimpancé. Los resultados experimentales se analizaron teniendo en cuenta estas dos hipótesis de organización del *locus*, pero privilegiando el arreglo de cinco genes (hipótesis B) debido a las limitaciones de secuencias que impone el arreglo de seis genes (hipótesis A). A su vez, en esta figura se señala la región cuyo análisis resolvería el dilema en cuestión.

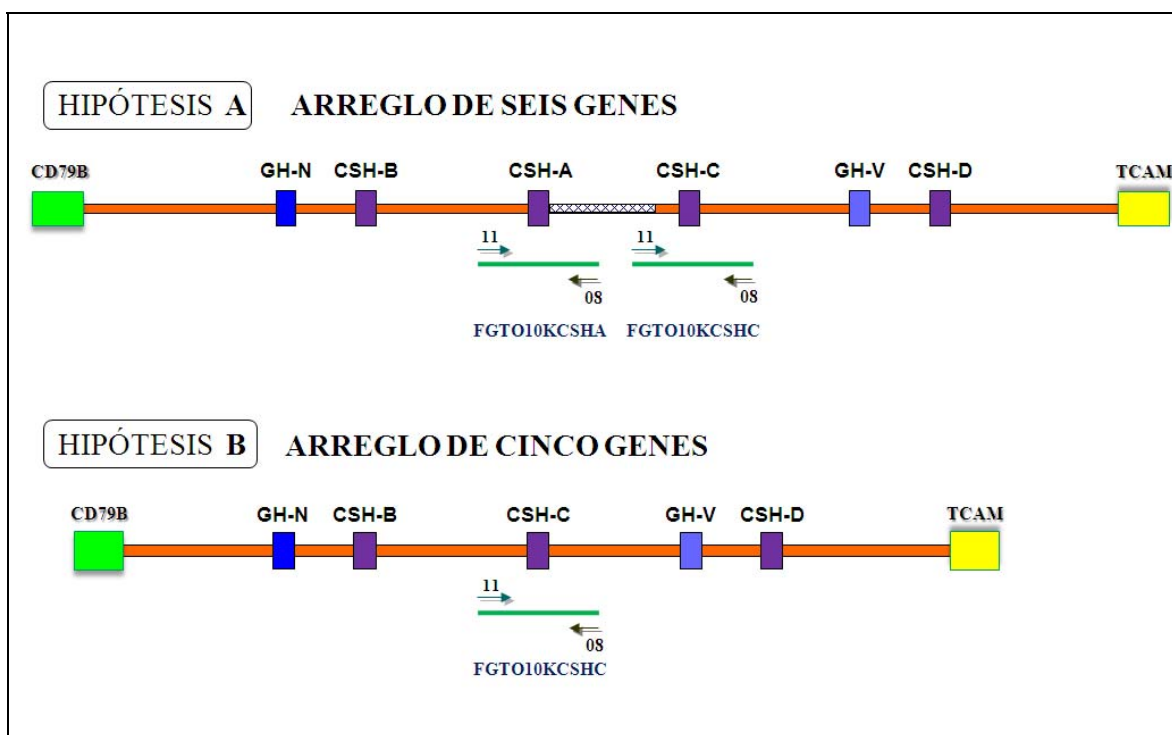


Figura 17. Arreglos alternativos de genes propuestos para el *locus* GH del chimpancé. Como hipótesis A se propuso un arreglo de seis genes y como hipótesis B uno de cinco. Se representan con líneas horizontales los probables productos amplificados a obtener de la región crítica para discernir los modelos. Con flechas en sus extremos se indica el par de iniciadores que amplifica esta región, que por comodidad se decidió mantener con la clave de identificación (FGTO10K), misma con la que se manejó durante la realización de este trabajo. En los extremos del *locus* se representan los genes CD79B y TCAM que flanquean al mismo.

En la figura 18 se muestra un gel típico en el que se tamizan las clonas recombinantes obtenidas para el FGTO10K por retraso del ADN plasmídico extraído a partir de estas. En

los carriles del 3 al 7 se observa un retraso en la migración electroforética que corresponde con las clonas que incorporaron el fragmento de interés (FGTO10K).

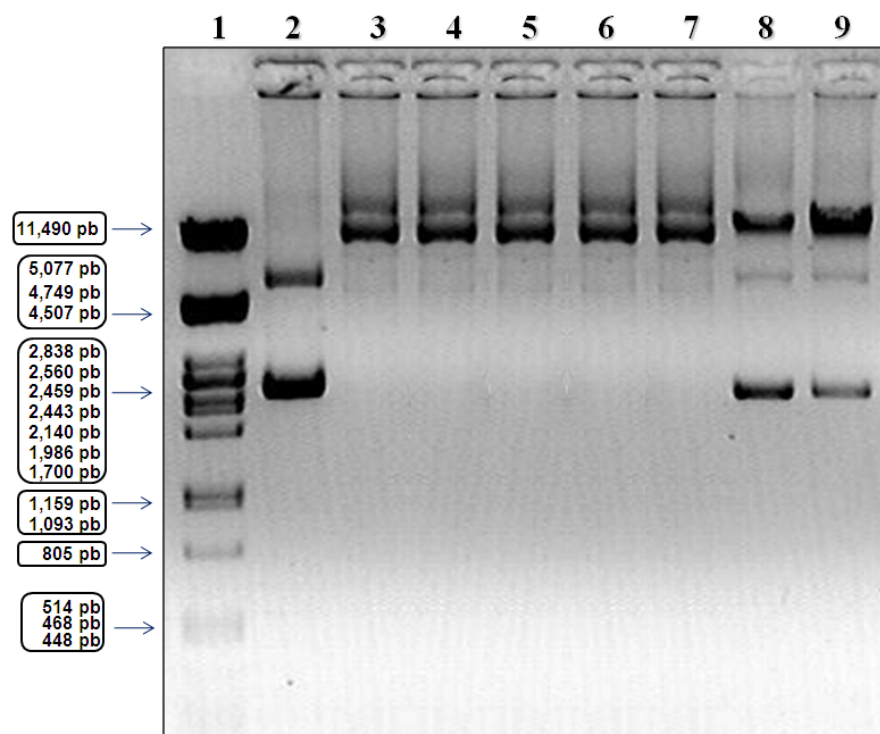


Figura 18. Tamizaje de las clonas recombinantes obtenidas para el FGTO10K. Electroforesis en gel de agarosa al 1%. En el carril 1 como marcador de peso molecular λ , en el carril 2 como control negativo el ADN del vector sin inserto, en los carriles del 3 al 9 se analizó el ADN plasmídico extraído a partir de las clonas recombinantes.

El producto amplificado obtenido para el FGTO10K fue caracterizado por cortes con enzimas de restricción, cotejando contra la predicción a partir de su secuencia nucleotídica. FGTO10K abarca regiones de dos intergénicas largas consecutivas y el gen incluido entre ellas, que para el caso de la hipótesis A resultaría en dos fragmentos, uno de los cuales incluiría al gen CSH-A y el otro al gen CSH-C. Esta región es crítica en la diferenciación de los dos posibles arreglos de genes, pues en el caso de la hipótesis B únicamente se amplificaría el fragmento que incluye al gen CSH-C. Para la hipótesis A se obtendría la combinación de los patrones de restricción de ambos fragmentos, mientras que para la hipótesis B se obtendría sólo el patrón correspondiente al fragmento que incluye al gen CSH-C, coincidiendo este último con el resultado experimental obtenido mostrado en la figura 19.

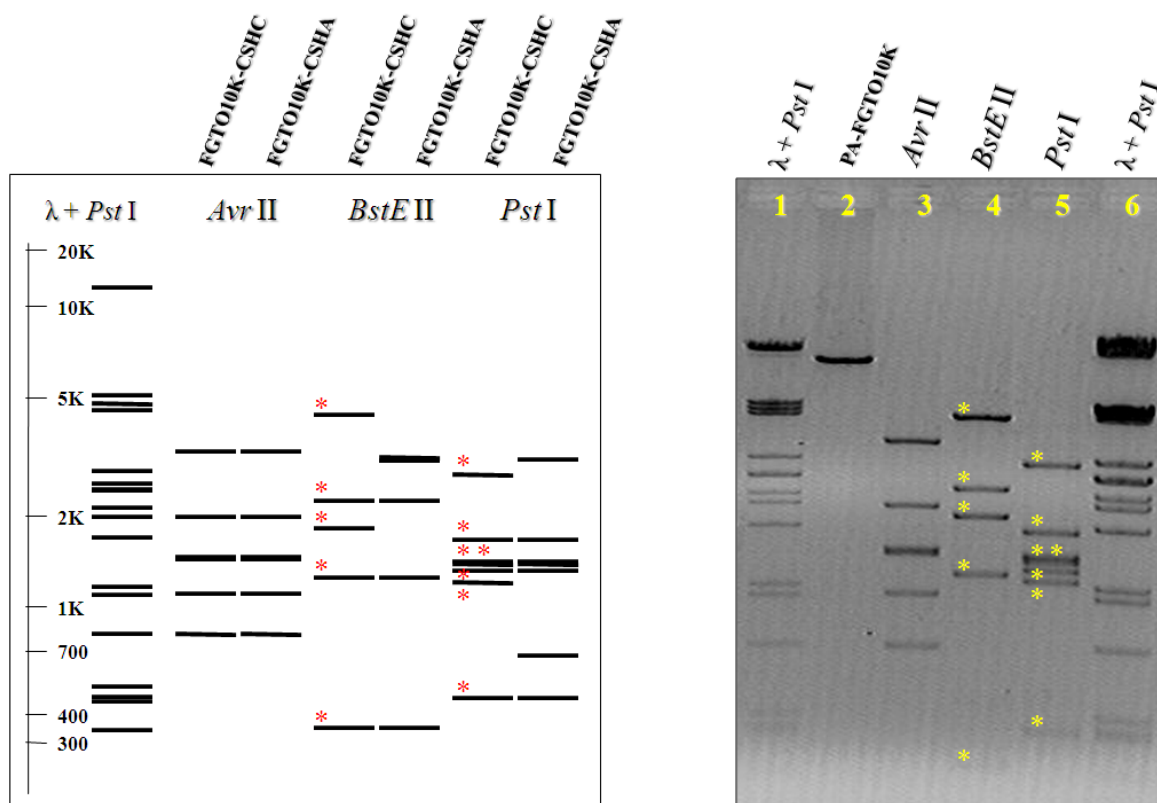


Figura 19. Caracterización del producto amplificado obtenido para FGTO10K. A la izquierda se muestran los patrones de digestión predichos para los dos fragmentos de tamaño similar que se amplificarían según el arreglo de seis genes (para el arreglo de cinco genes solo se amplificaría el fragmento que incluye al gen CSH-C y para el de seis se obtendría la suma de ambos). Con asteriscos se representan las bandas predichas *in silico* que están en concordancia con los resultados experimentales del gel de agarosa al 1% de la derecha, en el cual se verificaron los productos de digestión. En los carril 1 y 6 marcador de peso molecular λ , carril 2 corresponde al producto amplificado sin digerir, los carriles 3, 4 y 5 corresponden al producto amplificado digerido con las enzimas de restricción *AvrII*, *BstEII* y *PstI*, respectivamente.

5.6. Un nuevo enfoque para verificar el ensamblaje del *locus*.

Un segundo enfoque abordado para verificar el modelo de cinco genes y descartar el de seis, fue clonar los productos amplificados y caracterizar los plásmidos recombinantes con enzimas de restricción para obtener por separado los patrones. Al predecir el patrón de corte para cada clona se tuvo en cuenta la orientación del inserto dentro del vector, para evitar errores de interpretación. En las figuras 20, 21 y 22 se muestran las caracterizaciones de las clonas recombinantes obtenidas para el FGTO10K, con las enzimas de restricción *AvrII*, *BstEII* y *PstI*. Los análisis con las enzimas de restricción para todas las clonas recombinantes concuerdan nuevamente con la hipótesis B de cinco genes.

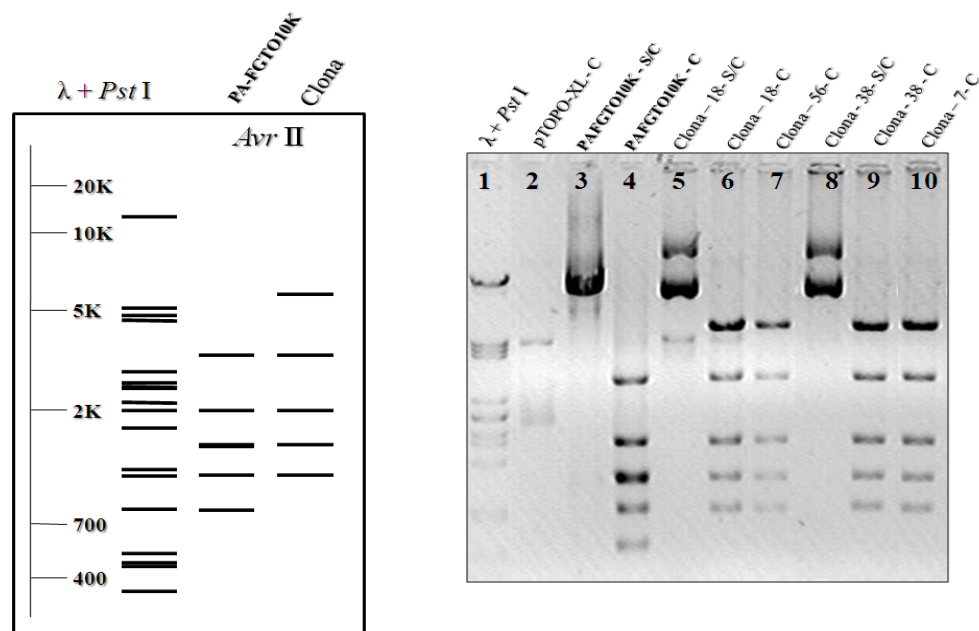


Figura 20. Caracterización con *AvrII* de las clonas recombinantes obtenidas para el FGTO10K. A la izquierda se muestran los patrones de digestión predichos para el producto amplificado del FGTO10K (PA-FGTO10K) y para la clona portadora de éste. A la derecha se muestra un gel de agarosa al 1% en el cual se verificaron los productos de digestión. En el carril 1 marcador de peso molecular λ; carril 2 corresponde al vector sin inserto digerido con *AvrII* (no presenta sitios de corte para esta enzima por lo cual permanece sin cortar); carril 3 corresponde al producto amplificado sin cortar; carril 4 al producto amplificado digerido; carriles 5 y 8 a clonas recombinantes sin digerir; y carriles 6, 7, 9 y 10 a clonas recombinantes digeridas.

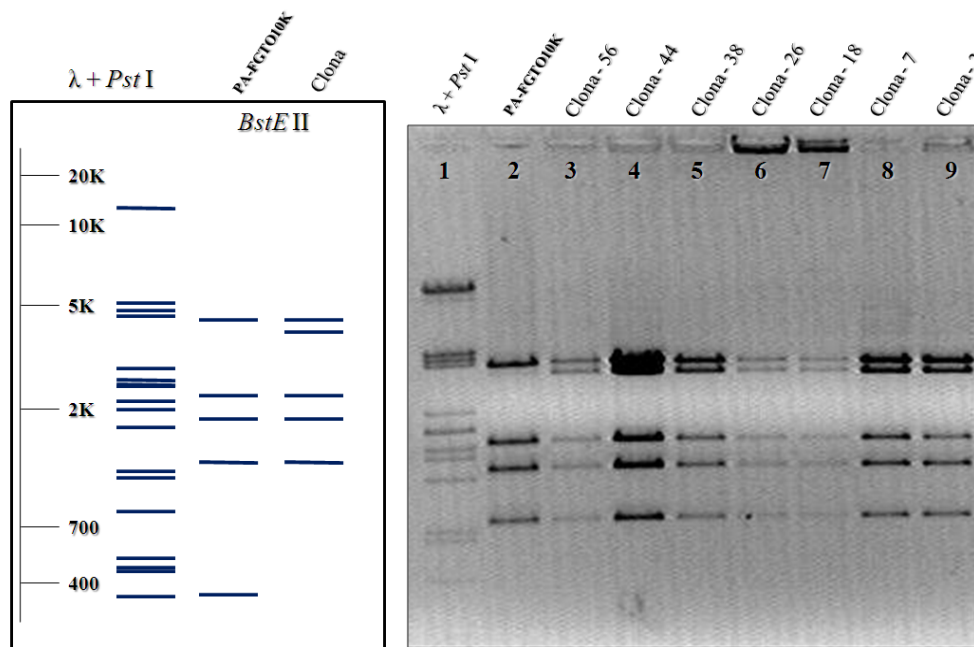


Figura 21. Caracterización con *BstEII* de las clonas recombinantes obtenidas para el FGTO10K. A la izquierda se muestran los patrones de digestión predichos para el producto amplificado del FGTO10K (PA-FGTO10K) y para la clona portadora de éste. A la derecha se muestra un gel de agarosa al 1% en el cual se verificaron los productos de digestión. En el carril 1 marcador de peso molecular λ, carril 2 corresponde al producto amplificado digerido y carriles del 3 al 9 a clonas recombinantes digeridas.

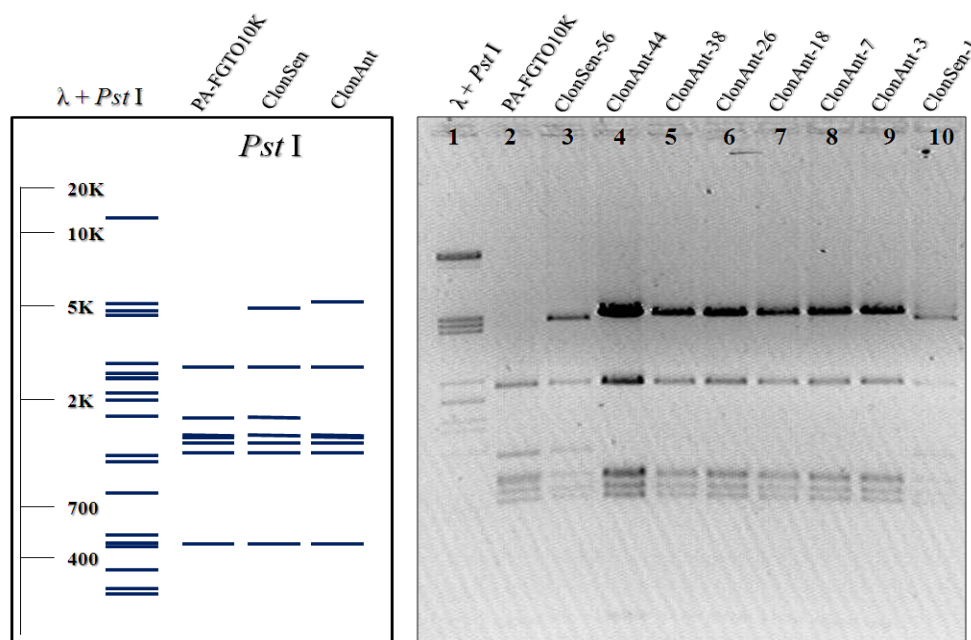


Figura 22. Caracterización con *Pst*I de las clonas recombinantes obtenidas para el FGTO10K. A la izquierda se muestran los patrones de digestión predichos para el producto amplificado del FGTO10K (PA-FGTO10K), la clona portadora de éste en orientación sentido (ClonSen) y en orientación antisentido (ClonAnt). A la derecha se muestra un gel de agarosa al 1% en el cual se verificaron los productos de digestión. En el carril 1 marcador de peso molecular λ , carril 2 corresponde al producto amplificado digerido, carriles del 3 al 10 a clonas recombinantes digeridas.

Una vez corroborado que las clonas recombinantes portadoras del FGTO10K contenían exclusivamente al gen CSH-C, el mismo fue amplificado a partir de estas, clonado y secuenciado.

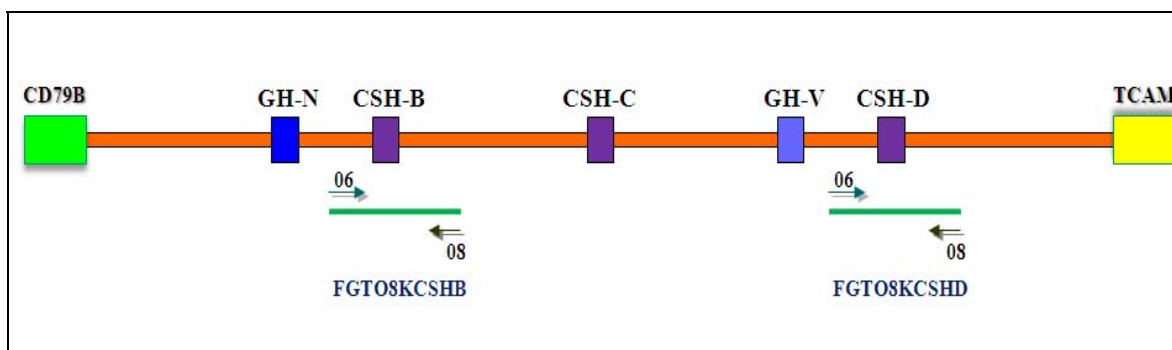


Figura 23. Representación de las regiones amplificadas con el par de iniciadores 06-08. Basado en el arreglo de cinco genes se representan con líneas horizontales los probables productos amplificados a obtener y en sus extremos (con flechas) el par de iniciadores que amplifica estos fragmentos.

De forma similar se procedió con FGTO8K obtenido con el par de iniciadores 06-08 que amplifican una región de intergénica corta, seguida de una región de intergénica larga (o la

región equivalente posterior al último miembro del *locus*) y el gen incluido entre ellas (ver la figura 23 para mayor claridad). Con esta amplificación se esperaba obtener dos fragmentos, uno de los cuales incluiría al gen CSH-B y el otro al gen CSH-D. Los productos amplificados fueron clonados y las clonas tamizadas mediante PCR en colonia con los iniciadores empleados para amplificar la unidad transcripcional de los genes GH/CSH. Posteriormente los productos amplificados fueron caracterizados mediante cortes con la enzima de restricción *StuI* para distinguir CSH-B de CSH-D (figura 24).

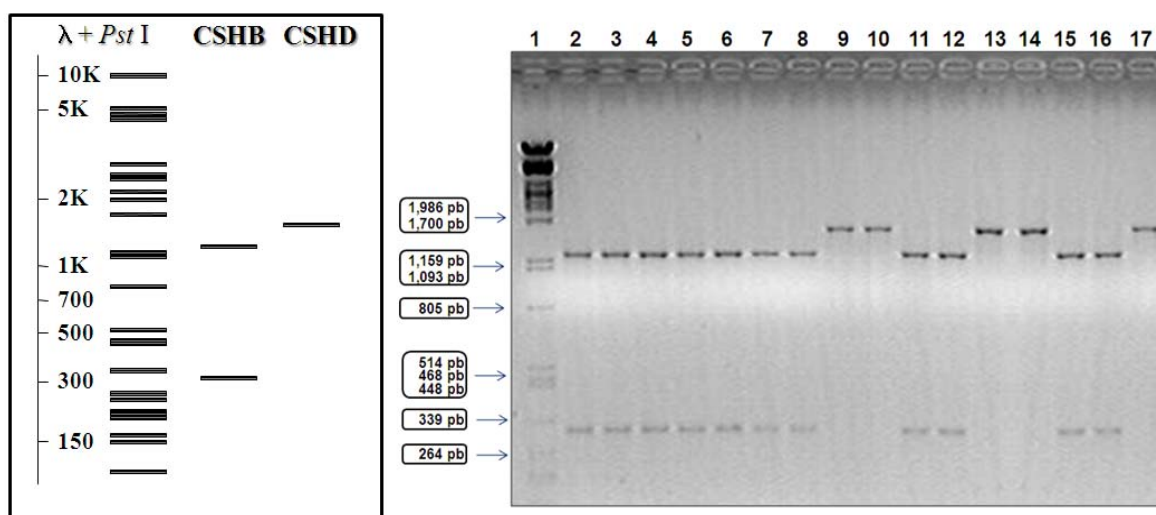


Figura 24. Caracterización de los productos amplificados por digestión con *StuI*. A la izquierda se muestra los patrones de digestión esperados en base a las secuencias. A la derecha se muestra un gel de agarosa al 2% en el cual se verificaron los productos de digestión. En el carril 1 marcador de peso molecular λ . Los carriles del 2 al 17 muestran los patrones obtenidos al digerir con *StuI* los productos amplificados a partir de las clonas portadoras del FGTO8K.

Como se observa en el gel los productos amplificados de los carriles 9, 10, 13, 14 y 17, exhiben un patrón de corte que permite identificarlos como CSH-D, mientras que el resto de los amplicones digeridos con *StuI* coinciden con los resultados esperados para el gen CSH-B. Una vez identificados estos dos genes por caracterización enzimática, fueron secuenciados.

En la figura 25 se muestra un gel de agarosa al 1% en el que se verifica la integridad del ADN plasmídico extraído a partir de las clonas recombinantes obtenidas para el FGTO8K. En los carriles del 2 al 7 se observa un retraso en la migración electroforética que se corresponde con las clonas que incorporaron dicho fragmento.

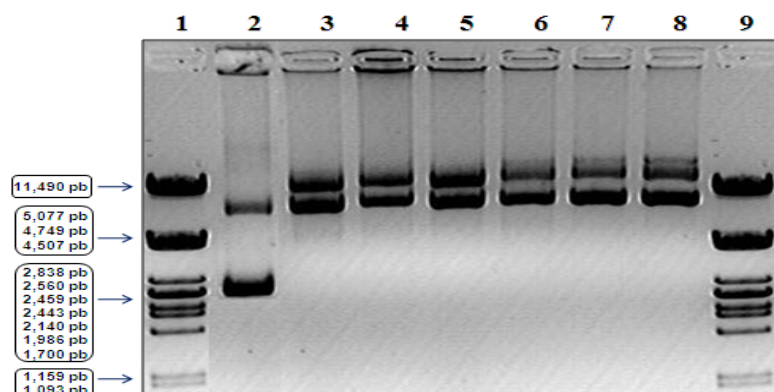


Figura 25. Tamizaje de las clonas recombinantes obtenidas para el FGTO8K. Electroforesis en gel de agarosa al 1%. En el carril 1 y 9 corresponden al marcador de peso molecular. En el carril 2 se muestra el ADN plasmídico del vector sin el inserto. En los carriles del 3 al 8 se verificó el ADN plasmídico extraído a partir de las clonas recombinantes.

La caracterización de dos de las clonas recombinantes obtenidas para el FGTO8K que son representativas de las regiones amplificadas y que corroboran el ensamblaje del *locus* GH del chimpancé realizado en este trabajo, se muestran en la figura 26. En todos los carriles de la misma se obtuvo correspondencia entre las predicciones y los resultados, sin bandas de más o de menos.

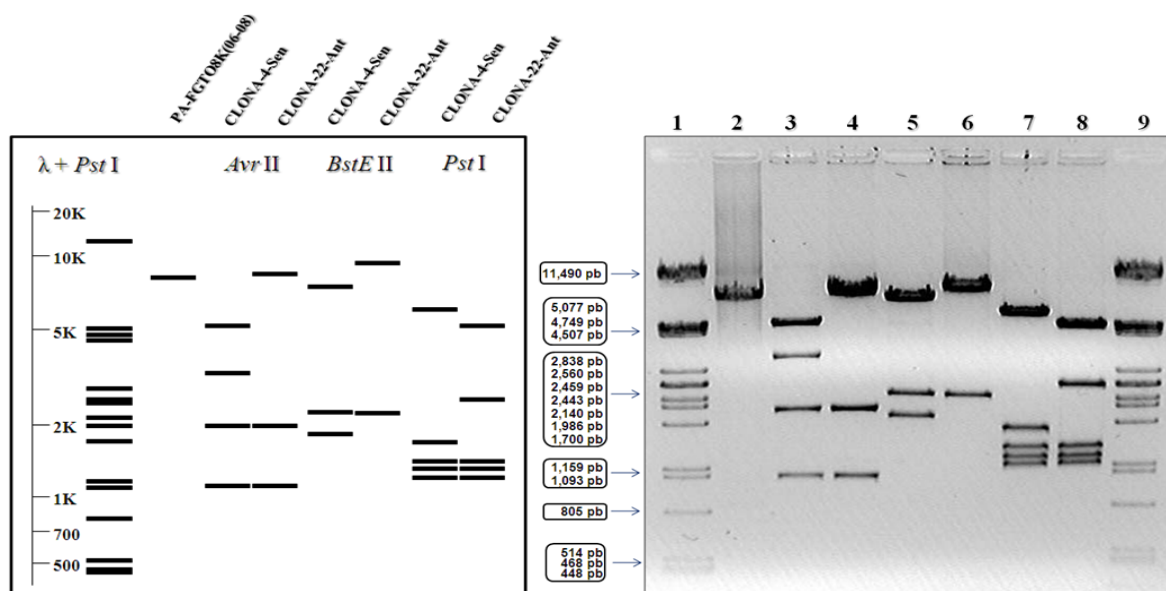


Figura 26. Caracterización de las clonas obtenidas para FGTO8K. A la izquierda se muestran los patrones de digestión predichos para los dos fragmentos de tamaño similar que se amplificarían, según el ensamblaje del *locus* GH del chimpancé propuesto en este trabajo, uno de los cuales incluye al gen CSH-B y el otro al gen CSH-D. A la derecha se muestra un gel de agarosa al 1% en el cual se verificaron los productos de digestión. En los carriles 1 y 9 marcador de peso molecular λ ; carril 2 corresponde al producto amplificado sin digerir; carriles 3, 5 y 7 corresponden a una de las clonas recombinantes (orientación sentido) representativa de la región amplificada que incluye al gen CSH-B, digerida con las enzimas *AvrII*, *BstEII* y *PstI*, respectivamente; los carriles 4, 6 y 8 revelan una de las clonas recombinantes (orientación antisentido) representativa de la región amplificada que incluye al gen CSH-D digerida con las mismas tres enzimas.

5.7. Confirmación del arreglo de cinco genes revela la organización del *locus* GH en el chimpancé.

Otro enfoque importante para la diferenciación de ambos arreglos de genes fue la amplificación de intergénicas largas con flancos en genes (IGLsP1) y posterior caracterización de los productos amplificados. Como se muestra en la figura 27, para el caso de la hipótesis A se esperaría la amplificación de cuatro regiones correspondientes a las intergénicas largas y la región posterior al último miembro del *locus*, mientras que para la hipótesis B sólo se amplificarían tres regiones: dos de ellas correspondientes a intergénicas largas (IGL1P1 e IGL2P1) y la tercera a la región posterior al último gen del *locus* (POSTP1). El producto amplificado IGLsP1 obtenido, que representa una mezcla de los tres fragmentos mencionados, fue caracterizado con las enzimas de restricción *AvrII*, *BstEII* y *PstI*, cotejando los resultados experimentales comparados con su respectiva predicción en base a secuencias nucleotídicas para la mezcla de los tres fragmentos derivados de la hipótesis B.

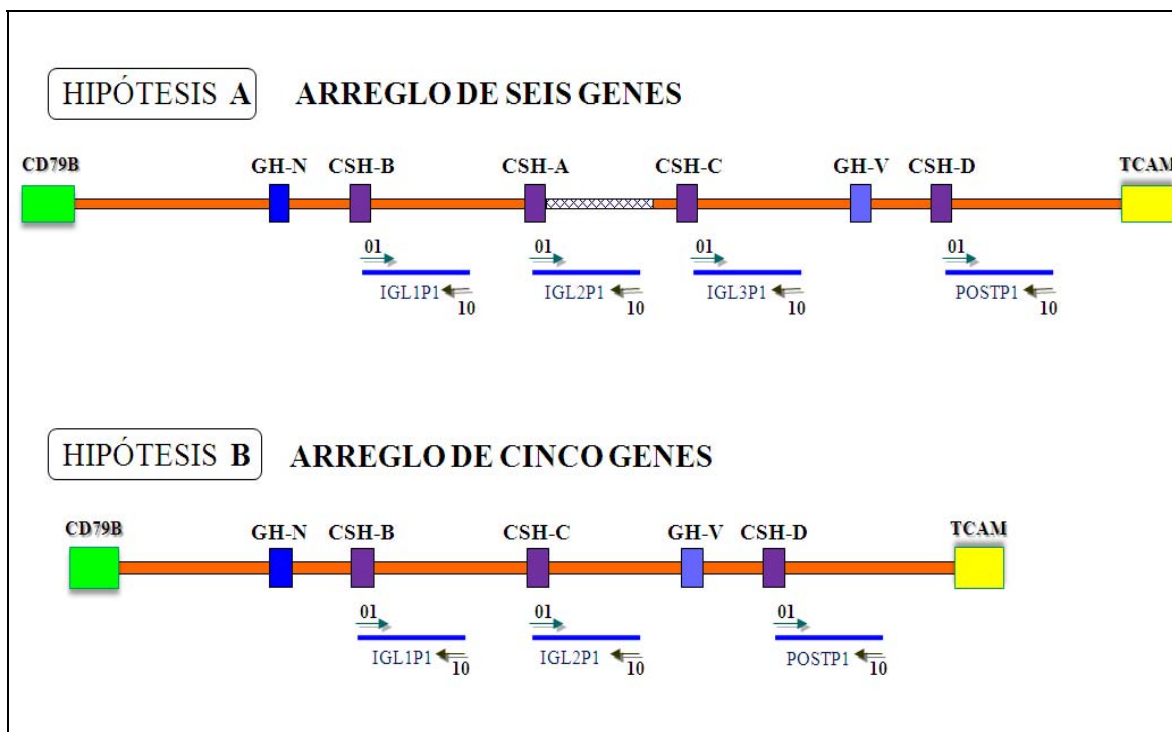


Figura 27. Representación de las regiones amplificadas con el par de iniciadores 01-10 en los dos arreglos de genes propuestos para el *locus* GH del chimpancé. Se representan con líneas horizontales los probables productos amplificados a obtener según cada hipótesis y con flechas en sus extremos el par de iniciadores que amplifica estos fragmentos.

En la parte superior de la figura 28 se muestra cómo se realizó la predicción de los patrones de cortes primero para cada fragmento en lo individual y luego para la suma de estos. Esta figura constituye el ejemplo representativo del procedimiento seguido en aquellos casos en los que se amplificó más de un fragmento en una misma reacción de PCR.

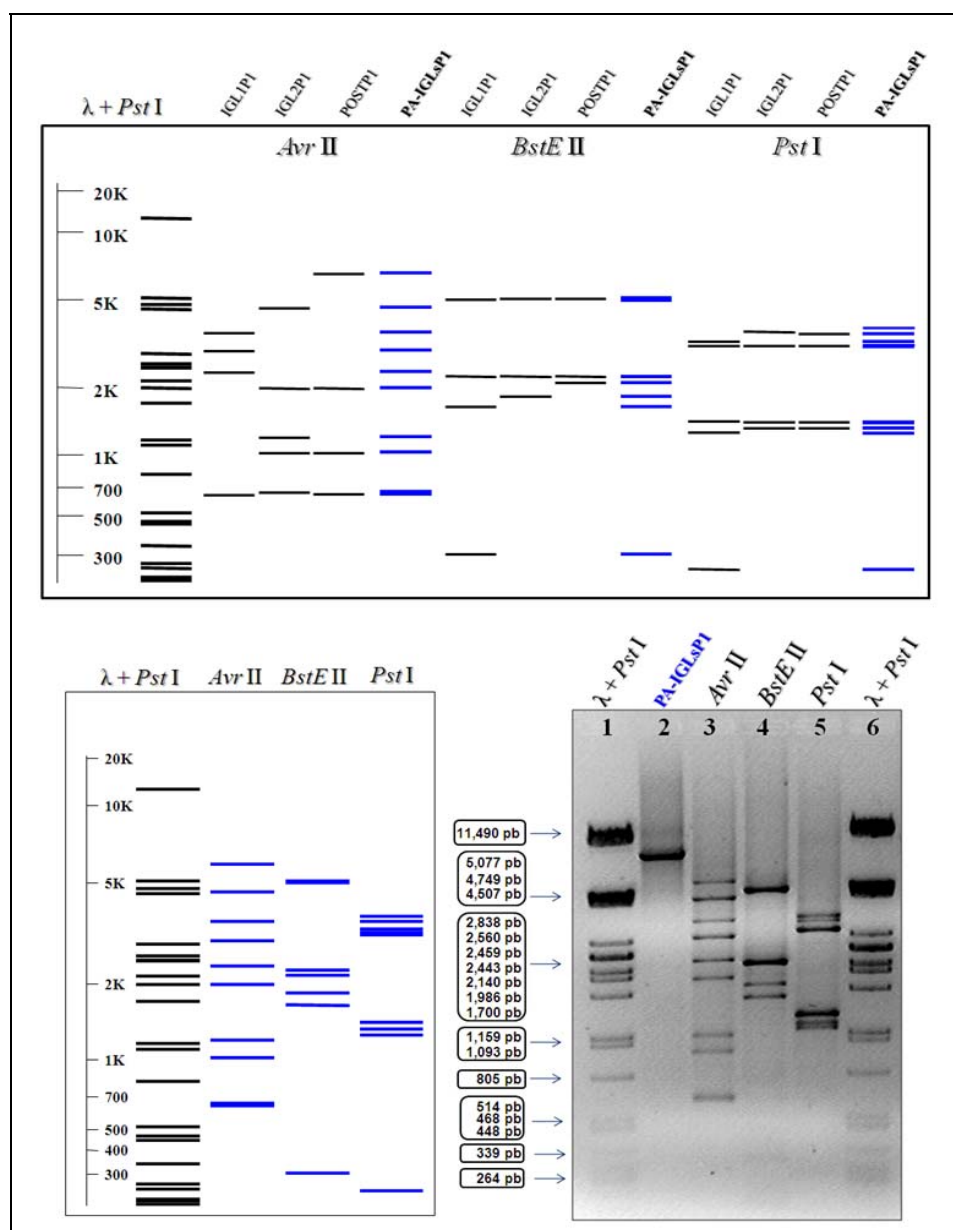


Figura 28. Caracterización del producto amplificado de las intergénicas largas con flancos en genes. En la parte superior se muestra cómo se realizó la predicción de los patrones de cortes para la mezcla de fragmentos. En la parte inferior izquierda se muestran los patrones de digestión predichos para la mezcla de fragmentos que se amplificarían, según la hipótesis B y a la derecha se muestra un gel de agarosa al 1% en el cual se verificaron los productos de digestión. En los carril 1 y 6 marcador de peso molecular; el carril 2 corresponde al producto amplificado sin digerir; los carriles 3, 4 y 5 corresponden al producto amplificado digerido con las enzimas de restricción *AvrII*, *BstEII* y *PstI*, respectivamente.

Como se puede observar en la foto del gel, presentó el patrón de corte esperado de la suma de cada fragmento con cada enzima de restricción. En el caso de que el *locus* se ajustara a la hipótesis A, este patrón se vería afectado con la presencia de varias bandas adicionales.

Adicionalmente se caracterizaron otras regiones amplificadas del *locus* que corroboran el ensamblaje del mismo (figura 29). En todos los casos analizados, los resultados esperados (predicción *in silico* en base a secuencia) para el ensamblaje del *locus* del chimpancé basado en la hipótesis B, coinciden con los resultados experimentales obtenidos.

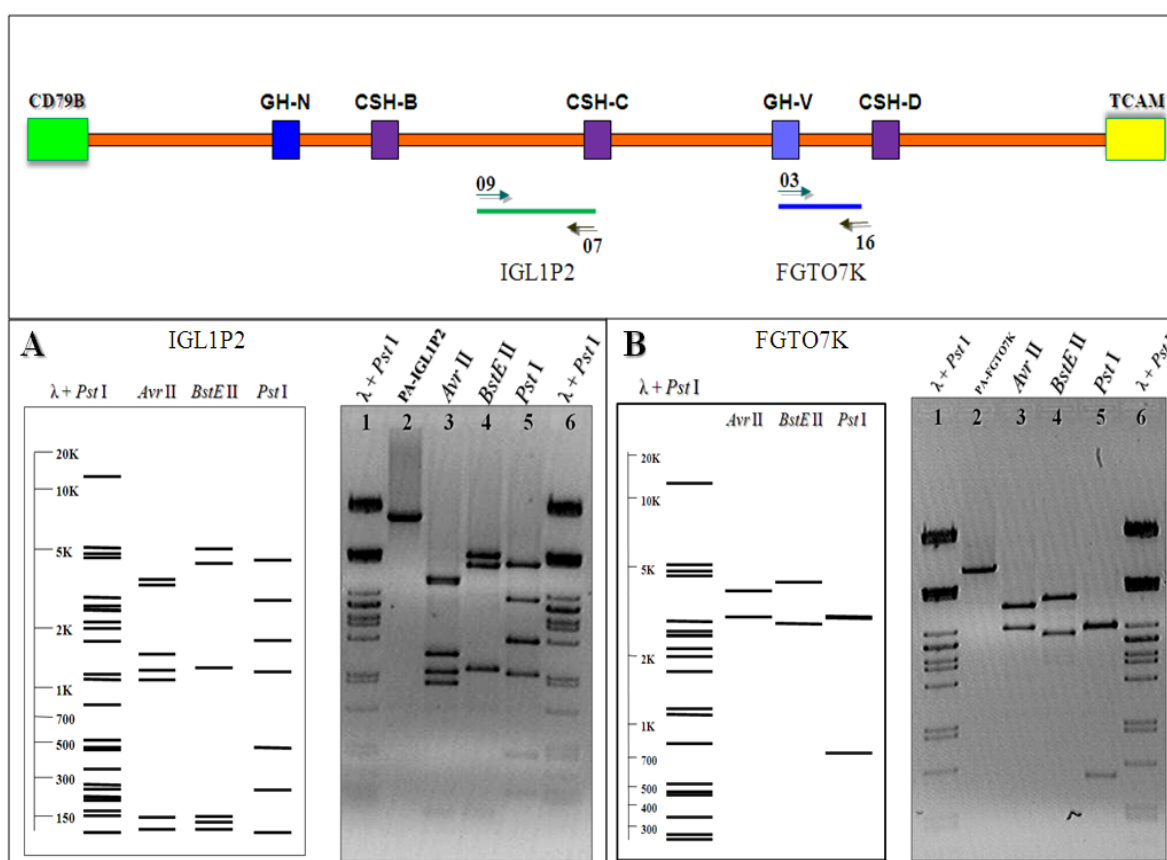


Figura 29. Caracterización de los amplicones obtenidos para otras regiones del *locus*. A la izquierda de ambos paneles se muestran los patrones de digestión predichos para IGL1P2 (Panel A) amplificado con los oligonucleótidos 09–07 que abarca una región de la primera intergénica larga y la mitad de la CSH-C, así como para FGTO7K (Panel B), que comprende la mitad del la GH-V y casi la totalidad de la intergénica corta 2, obtenido con los oligonucleótidos 03–16. A la derecha de ambos paneles se muestra el gel de agarosa al 1% en el cual se verificaron los productos de digestión. En los carril 1 y 6 marcador de peso molecular, carril 2 corresponde al producto amplificado sin digerir, los carriles 3, 4 y 5 corresponde al producto amplificado digerido con las enzimas de restricción *AvrII*, *BstEII* y *PstI*, respectivamente.

5.7. Ensamblaje final del *locus* GH.

Con las secuencias obtenidas en este trabajo, los datos de secuencias extraídos del GenBank y las caracterizaciones enzimáticas, fue posible determinar la organización genómica del *locus* GH del chimpancé, tal y como se muestra en la figura 30.

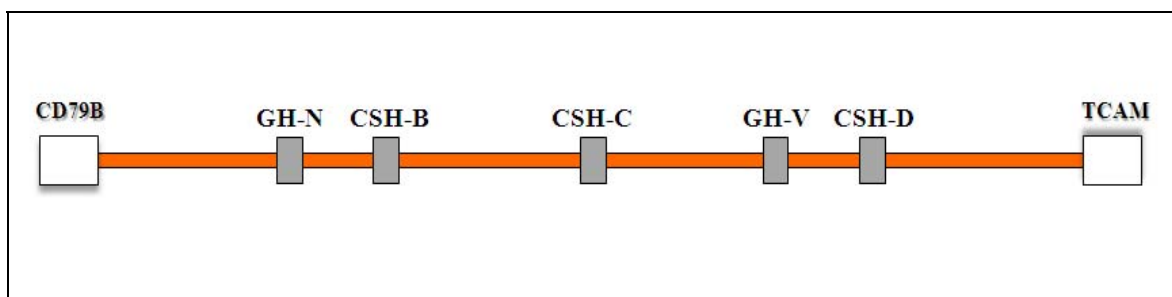


Figura 30. Organización del *locus* GH. Las cajas en gris indican los genes obtenidos, mientras que las vacías a los *loci* que flanquean a este *locus*. Las líneas sólidas representan a las regiones intergénicas obtenidas al complementar nuestras secuencias con las reportadas por el CSGCh.

5.8. Análisis de las proteínas.

Se obtuvieron cinco secuencias genómicas correspondiente a los genes que integran el *locus* GH del chimpancé, las cuales se utilizaron para predecir con la ayuda del programa Genescan las secuencias codificantes y sus secuencias aminoacídicas respectivas. En la figura 31 se muestra el alineamiento de las proteínas GH/CSH del chimpancé y del humano, usando como referencia a la secuencia del gen de la GH hipofisaria del humano (hGH-N). Para el pseudogen humano se predijo la proteína que este codificaría si no tuviese afectado ningún sitio de “splicing”.

Resultó de nuestro interés investigar si las diferencias aminoacídicas encontradas ocurrían en los sitios clave de unión al receptor. Adicionalmente se le prestó una especial atención a aquellos residuos que basado en datos experimentales se han reportado que juegan un papel determinante en la función, siendo estos en GHs: Ile4, Glu56, Arg64 e Ile179; mientras que en CSHs Val4, Asp56, Met64 y Met179 (Goffin *et al.*, 1996). También se tuvo en cuenta el residuo de fenilalanina de la posición 44 (Phe⁴⁴), ya que se ha sugerido que este aminoácido puede tener una importancia estructural particular en el sitio I de unión de la hormona con su receptor y estar relacionado con la actividad lactogénica de la misma (Peterson y Brooks, 1997, Duda y Brooks, 2003).

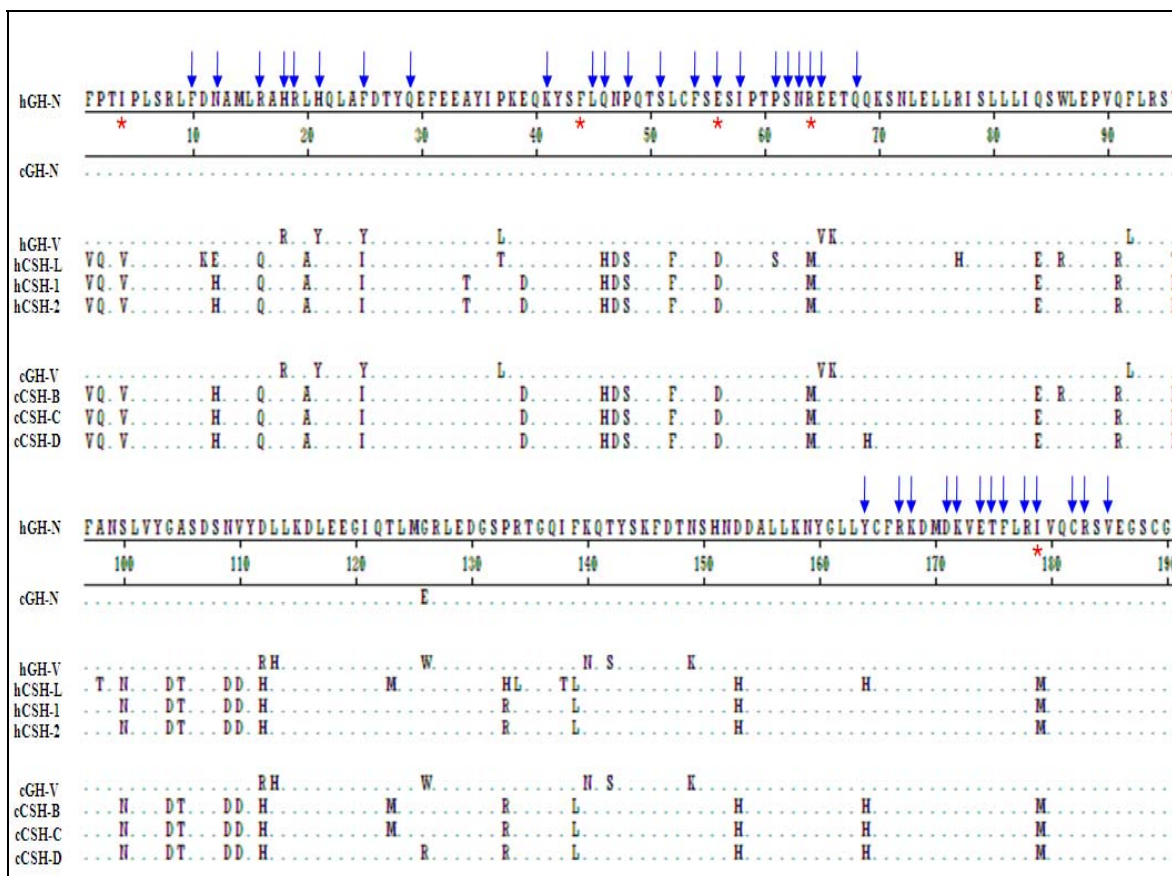


Figura 32. Alineamiento de las secuencias aminoacídicas. Se indican con flechas los aminoácidos claves en la unión al receptor reportados para la GH-N del humano (Liu *et al.*, 2001). Los puntos indican identidad con la secuencia de referencia. Los cambios aminoacídicos se representan por la letra correspondiente al aminoácido que difiere.

En el cuadro 8 se muestran los cambios que presentaron las proteínas deducidas a partir de las secuencias de los genes obtenidos en este trabajo, respecto a su contraparte en el humano (Chen *et al.*, 1989). La GH-N inferida en nuestro estudio mostró una sola sustitución aminoacídica con respecto a su contraparte en el humano y la proteína GH-V resultó ser idéntica en ambas especies. Las CSH-C y CSH-D del chimpancé con tres y cuatro cambios aminoacídicos respecto a sus ortólogas en el humano, respectivamente, tuvieron en común uno de estos cambios (sustitución de tirosina por histidina en la posición 164). La CSH-B del chimpancé con diez cambios aminoacídicos respecto a su contraparte en el humano (CSH-L) fue la que exhibió la mayor divergencia, algunas de estas sustituciones ocurren en sitios clave de unión al receptor.

Cuadro 8. Cambios aminoacídicos entre las proteínas inferidas en este trabajo y sus contrapartes en el humano.

Parejas génicas	Cambios	Posiciones	Sustituciones
cGH-N vs hGH-N	1	126	Gly x Glu
cCSH-B vs hCSH-L*	10	11	Lys x Asp
		12	Glu x His
		37	Thr x Pro
		61	Ser x Pro
		77	His x Arg
		96	Thr x Met
		133	His x Arg
		134	Leu x Arg
		138	Thr x Ile
cCSH-C vs hCSH-1	3	34	Thr x Ala
		123	Thr x Met
		164	Tyr x His
cGH-V vs hGH-V	0	-	-
cCSH-D vs hCSH-2	4	34	Thr x Ala
		69	Gln x His
		126	Gly x Arg
		164	Tyr x His

* Se ignoró la mutación clave de este pseudogen para predecir su proteína.

5.9. Comparaciones de los genes placentarios de los *loci* versus los hipofisiarios.

En la figura 33 se muestra un análisis comparativo de los genes GH-V/CSH de humano y chimpancé en relación a sus respectivas GH-N. En ambas especies todos los genes exhiben una similitud superior al 92% respecto a la GH-N, excepto en el caso del pseudogen humano que exhibe un valor ligeramente menor.

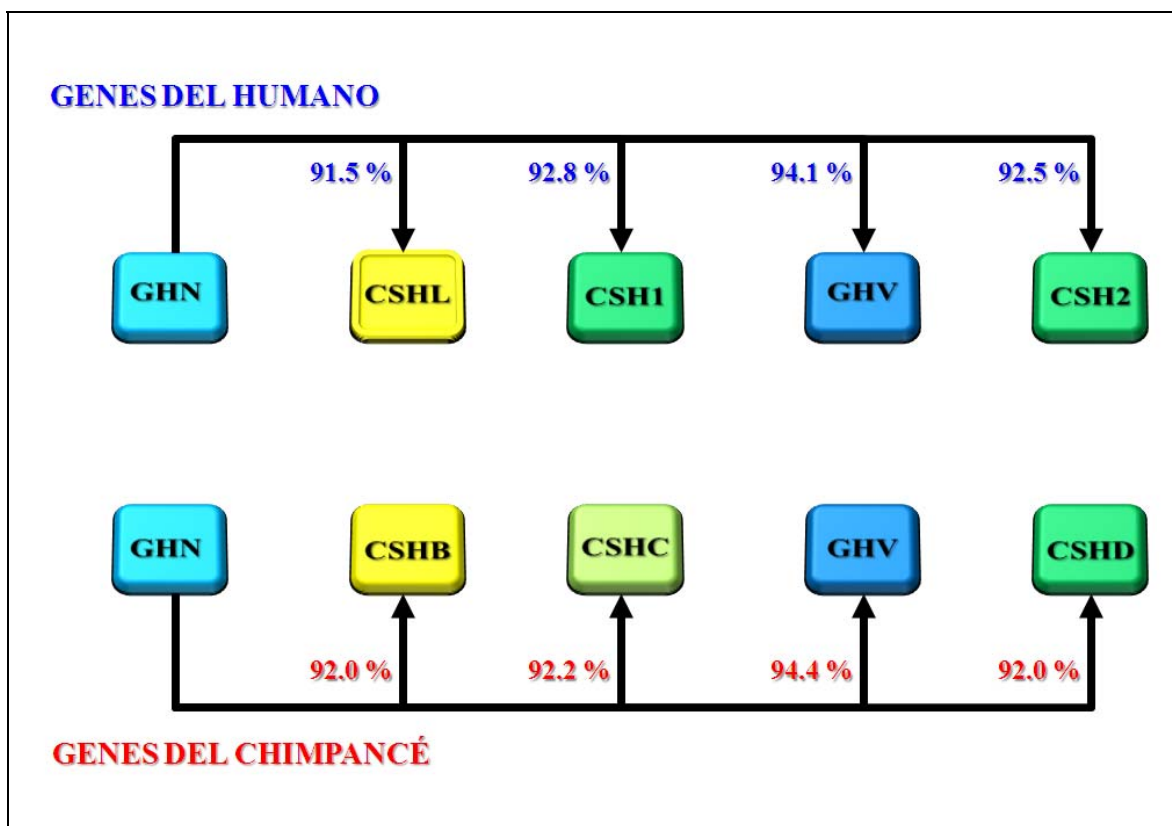


Figura 33. Análisis comparativo de los genes GH/CSH parálogos, con respecto a las GH-N. Se muestran en porcentaje los valores de similitud entre los genes comparados obtenidos del alineamiento de sus secuencias mediante el método de ClustalW.

En la figura 34 se muestra un análisis comparativo entre los genes GH/CSH ortólogos de humano y chimpancé, indicándose los respectivos porcentajes de similitud obtenidos del alineamiento de sus secuencias nucleotídicas. Todos los genes del chimpancé exhiben una similitud cercana o superior al 98% respecto a su ortólogo en el humano, excepto CSH-B cuyo valor fue 96%. Interesantemente este gen mostró mayor similitud con el gen CSH-1 (98.3%). Igualmente llama la atención que el gen CSH-D de chimpancé exhiba una mayor similitud con CSH-1 (98.2%) que con su probable ortólogo.

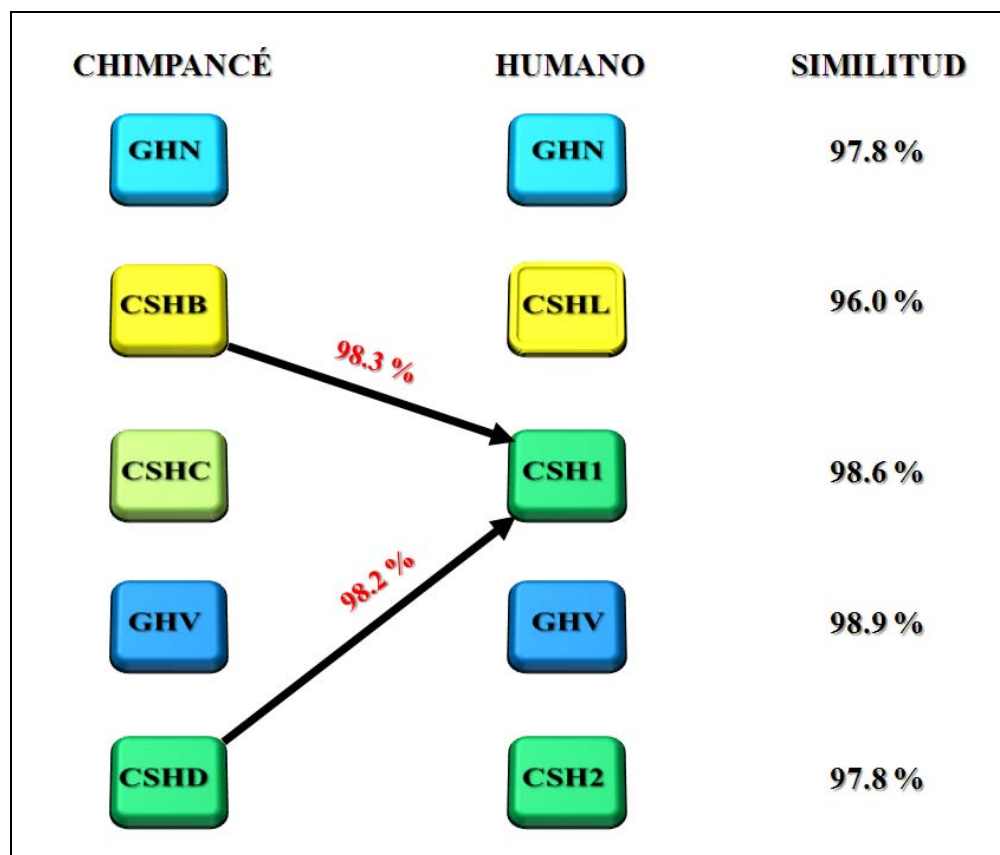


Figura 34. Análisis comparativo entre genes ortólogos. Se muestran en porcentaje los valores de similitud entre los genes que se presumen ortólogos, obtenidos del alineamiento de sus secuencias mediante el método de ClustalW.

5.10. Comparaciones de las proteínas maduras GH/CSH de humano y chimpancé.

La figura 35 muestra un análisis comparativo de las proteínas maduras de humano y chimpancé respecto a la GH-N, donde se indican los respectivos porcentajes de similitud obtenidos del alineamiento de sus secuencias aminoacídicas.

Las CSHs de humano y chimpancé exhiben una similitud del 83.8 al 85.3% respecto a la GH-N, excepto la secuencia predicha para el pseudogen humano, cuyo valor desciende hasta 81.2%. Por otro lado, tanto la GH-V de humano como la de chimpancé, por ser idénticas, mostraron igual porcentaje de similitud (93.2%) respecto a la GH-N.

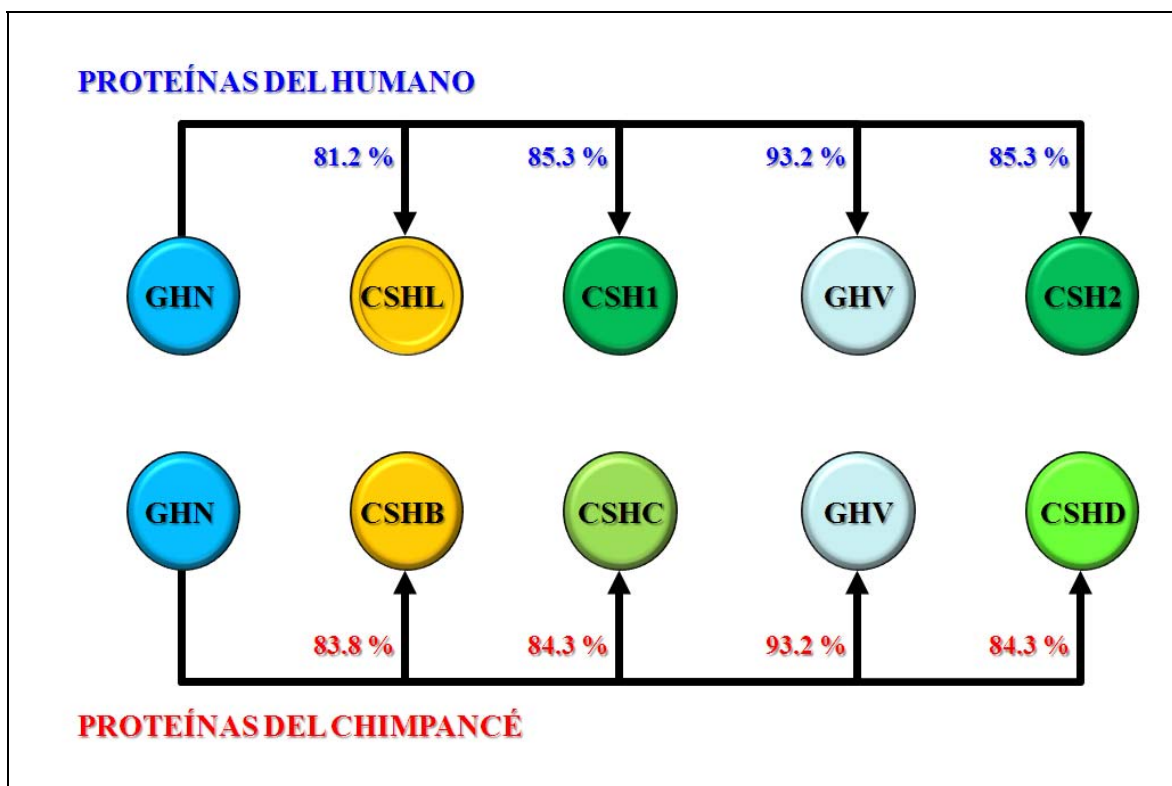


Figura 35. Análisis comparativo de las proteínas maduras de humano y chimpancé respecto a GH-N. Se muestran en porcentaje los valores de similitud entre las secuencias de las proteínas maduras comparadas obtenidos del alineamiento de sus secuencias mediante el método de ClustalW. Las flechas indican similitudes entre no ortólogos.

En la figura 36 se muestra un análisis comparativo entre las proteínas ortólogas de humano y chimpancé, indicándose los respectivos porcentajes de similitud obtenidos del alineamiento de sus secuencias aminoacídicas. Todas las secuencias aminoacídicas del chimpancé exhiben una similitud arriba del 97.9% respecto a su ortólogo en el humano, excepto CSH-B cuyo valor fue 94.2%. Interesantemente, esta proteína mostró mayor similitud con CSH-1 y CSH-2 (97.9%). La secuencia aminoacídica de la GH-V de humano y chimpancé son idénticas. La GH-N de humano y chimpancé comparten un 99.5% de similitud en sus secuencias proteicas (un solo cambio aminoacídico).

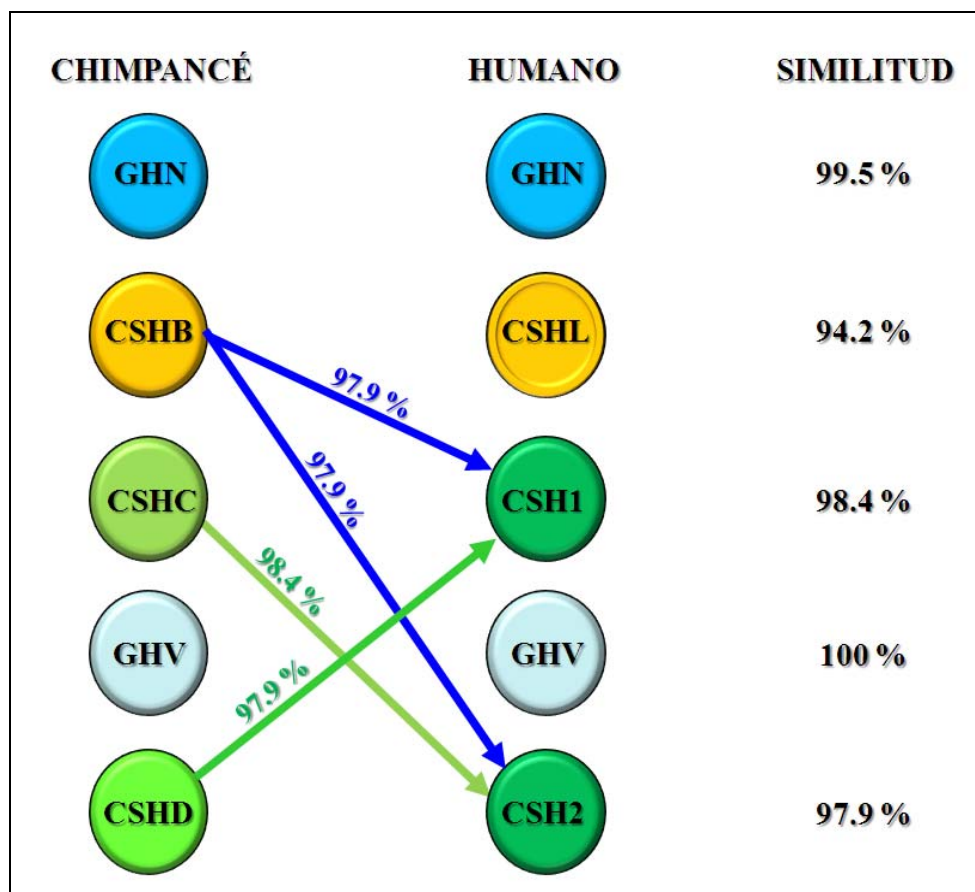


Figura 36. Análisis comparativo entre proteínas ortólogas. Se muestran en porcentaje los valores de similitud entre las secuencias de las proteínas maduras comparadas obtenidos del alineamiento de sus secuencias mediante el método de ClustalW. Además del valor para las parejas de hormonas codificadas por genes que se presumen ortólogos, también se hace esto para aquellas otras con alta similitud.

5.11. Análisis filogenético de los genes GH/CSH de humano y chimpancé.

En la figura 37 se muestra el árbol filogenético de los genes GH/CSH de humano y chimpancé, en cuya construcción se incluyó como grupo fuera la secuencia del cerdo. Sólo se observa clara ortología entre los genes GH-N y GH-V de humano y chimpancé, mientras que para las CSHs se observa un agrupamiento especie-específico y no se aprecia una clara ortología entre ellas. Adicionalmente se utilizaron otros métodos como máxima parsimonia y “maximum likelihood” para construir el árbol, pero siempre se obtuvo la misma topología (González *et al.*, 2006).

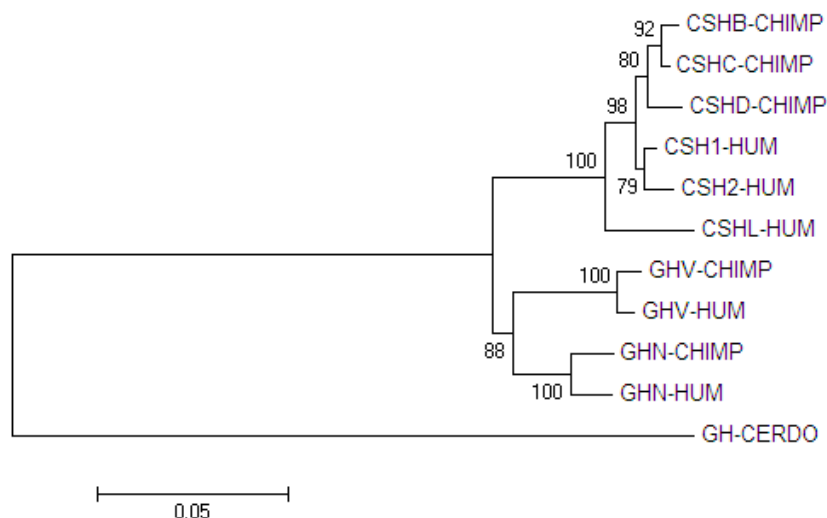


Figura 37. Árbol filogenético de los genes GH/CSH de humano y chimpancé. El árbol fue construido con el programa MEGA 3.0 por el método de NJ, utilizando las secuencias (desde el codón de inicio hasta el codón de paro) de los genes del chimpancé generadas en este estudio y las de humano y cerdo que se mencionan en el cuadro 6 de Material y Métodos.

5.12. Comparaciones de las regiones intergénicas de humano y chimpancé.

En la figura 38 se muestra un análisis comparativo entre las regiones intergénicas ortólogas de humano y chimpancé, indicándose los respectivos porcentajes de similitud obtenidos del alineamiento de sus secuencias nucleotídicas.

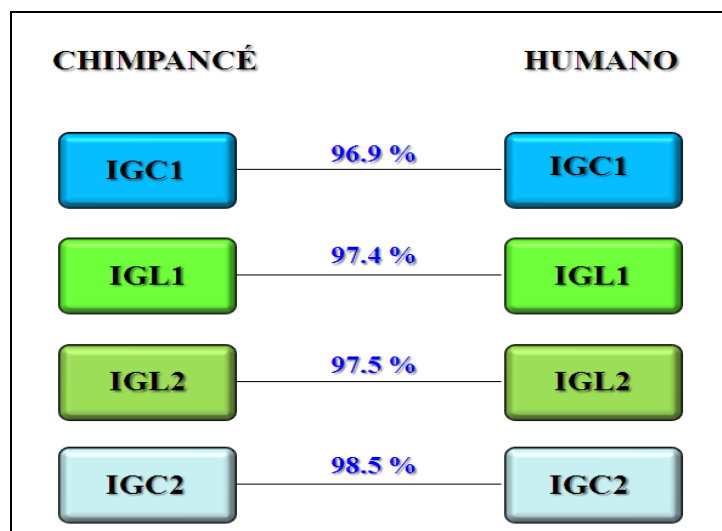


Figura 38. Análisis comparativo entre regiones intergénicas que se presumen ortólogas. Se muestran en porcentaje los valores de similitud entre las intergénicas comparadas obtenidos del alineamiento de sus secuencias mediante el método de ClustalW.

Todas las regiones intergénicas del chimpancé exhiben una similitud superior al 96.9% respecto a su ortólogo en el humano. Adicionalmente se determinó mediante el programa Repeatmasker (datos no mostrados) la presencia de elementos *Alu* en dichas intergénicas y se encontró que estos presentan la misma distribución que la previamente reportada para el *locus* humano (Chen *et al.*, 1989).

5.13. Un vistazo a reguladores de la transcripción.

5.13.1. El promotor proximal.

En la figura 39 se muestra el alineamiento de las secuencias correspondientes al promotor proximal de los genes GH/CSH de humano y chimpancé, indicándose los sitios de unión a factores transcripcionales descritos para el primero.

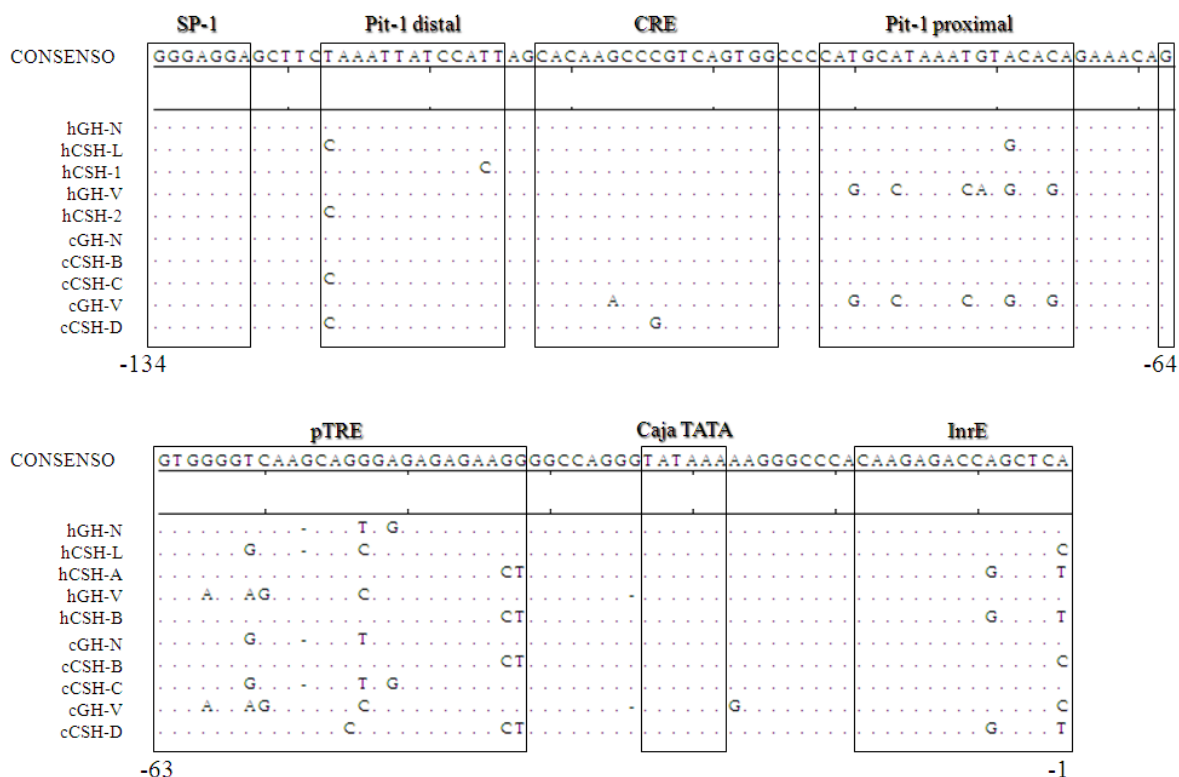


Figura 39. Alineamiento del promotor proximal. El alineamiento se hizo en el programa MEGALIGN, utilizando la secuencia del gen de la GH-N del humano como referencia, al momento de indicar los sitios de unión a los factores transcripcionales.

Estos sitios incluyen la caja TATA, los elementos Pit-1 distal y proximal (Theill y Karin, 1993; Krawczak *et al.*, 1999), un elemento de respuesta a AMPc (CRE) (Eberhard *et al.*,

1996), un elemento de respuesta a hormona tiroidea (TRE) (Glass et al., 1987), un sitio de unión al factor SP-1 (Lemaigre et al, 1989) y un sitio de unión del iniciador (InrE).

Los promotores proximales exhiben entre 96.6 y 98.6% de similitud entre sí y todos poseen los elementos mínimos indispensables para la transcripción basal. Sin embargo, se requieren ensayos de transfección frente algún gen reportero para evaluar el significado biológico de los cambios puntuales observados. En el caso de los elementos Pit-1, se puede observar que tanto el distal, como el proximal se encuentran altamente conservados respecto a la secuencia consenso, excepto para las GH-V de humano y chimpancé, cuyos promotores exhiben algunas variaciones, coincidiendo en ambos casos.

5.13.2. El inhibidor hipofisiario o elemento P.

Este elemento que inhibe la expresión de los genes placentarios en la hipófisis se encontró en todas las regiones intergénicas del chimpancé. En todos los casos, se localiza aproximadamente a 2.2 kpb río arriba de los genes, excepto para el gen CSH-D, que se ubica 300 pb aún más retirado. En la figura 40 se muestra el alineamiento de las 90 pb que comprenden a esta secuencia reguladora para humano y chimpancé. Todas las secuencias entre sí tienen una similitud arriba del 95.6%.

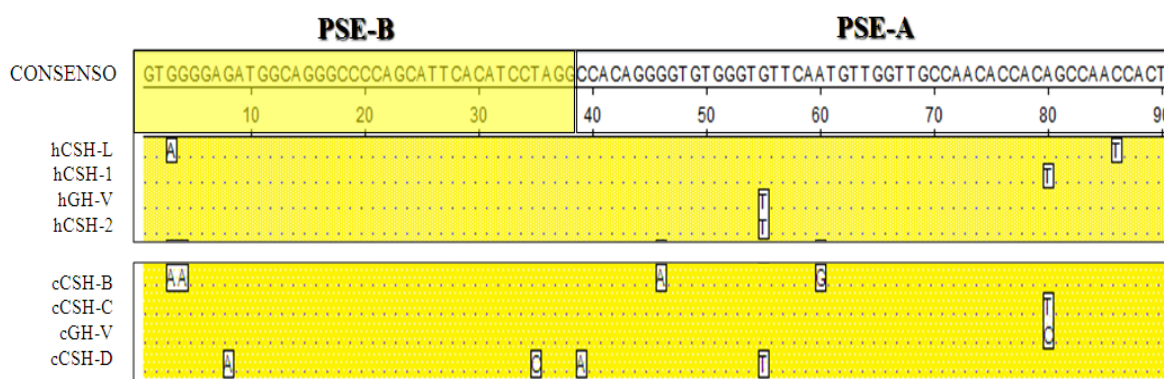


Figura 40. Alineamiento del inhibidor hipofisiario. Se obtuvo la secuencia de este elemento para cada uno de los genes mencionados y se alinearon con su contraparte en el humano.

5.13.3. El potenciador placentario o enhancer.

Este elemento que potencia la expresión de los genes CSH en la placenta, se encontró aproximadamente a 2.3 kpb después de estos genes en el chimpancé. En la figura 41 se muestra el alineamiento de las 280 pb que comprenden a esta secuencia reguladora en el chimpancé y en el humano.

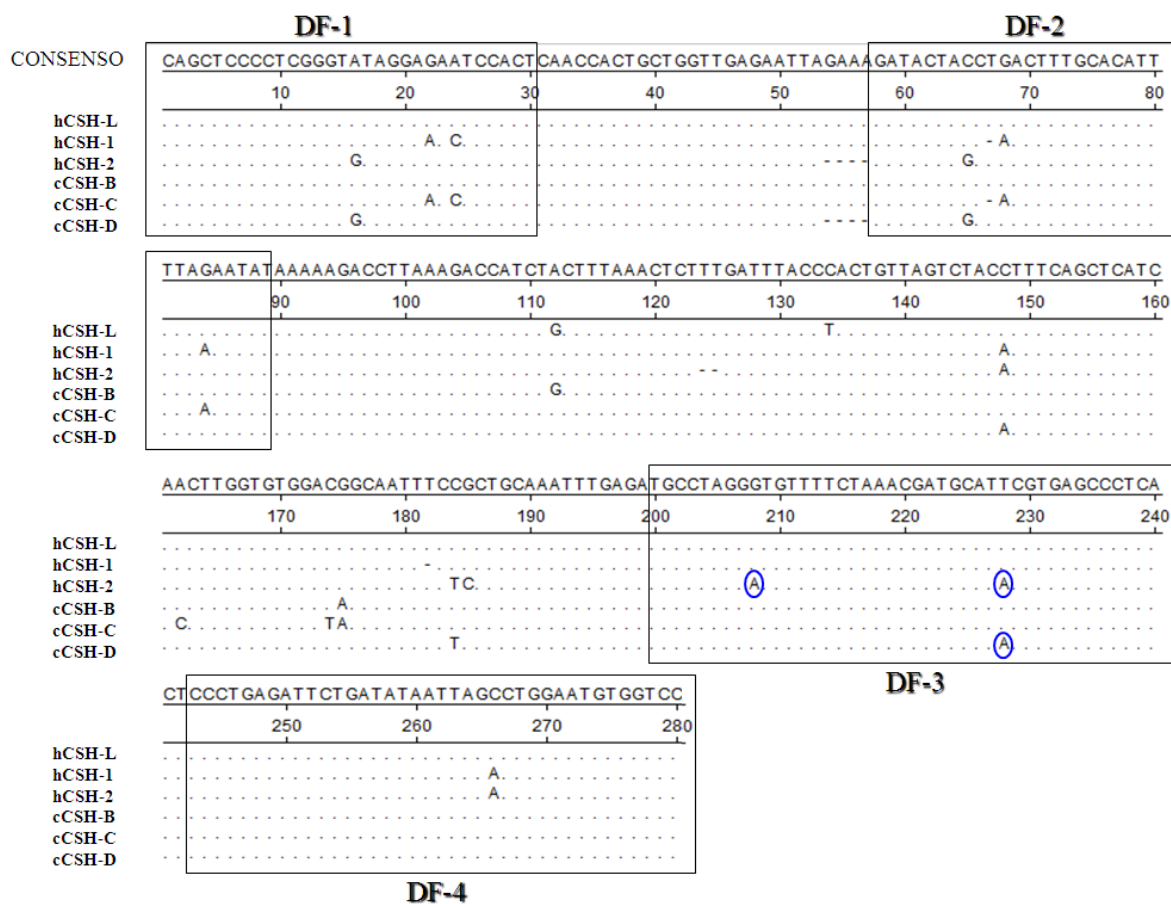
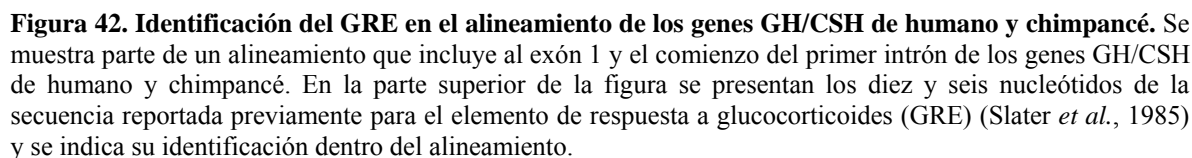


Figura 41. Alineamiento del potenciador placentario. Se obtuvieron las secuencias de los tres potenciadores del *locus* del chimpancé y se alinearon con sus contrapartes del humano. Con rectángulos se muestran los cuatro dominios DF del potenciador placentario. Encerrados en óvalos se indican los nucleótidos adenina que son blancos de mutaciones que reducen la actividad del potenciador.

Todas las secuencias entre sí exhibieron una similitud arriba del 97.9%. En el humano los potenciadores de los genes CSH-L y CSH-1 presentan una inserción de cuatro pares de bases río arriba del dominio DF2 que se ha demostrado que reducen la funcionalidad de estos (Jacquemin *et al.*, 1994). Esta misma inserción se encontró en los potenciadores de los genes CSH-B y CSH-C del chimpancé. En el dominio DF3 encerrados en un óvalo se

Se ha reportado la presencia de un elemento de respuesta a glucocorticoides (GRE) en el primer intrón del gen de la GH-N humana (Slater *et al.*, 1985). En la figura 42 se muestra parte de un alineamiento de las secuencias nucleotídicas de los genes GH/CSH de humano y chimpancé que muestra una región génica que abarca el exón 1 y parte del primer intrón. El GRE fue identificado en el alineamiento tomando como referencia el reporte previamente mencionado de su secuencia para la hGH. En este trabajo se determinó que dicho elemento se encuentra conservado en los genes GH-N y GH-V de chimpancé, notándose que también lo está en la GH-V humana.



CAPÍTULO VI: DISCUSION

El análisis de las secuencias de un genoma ofrece una amplia visión de su estructura, pero muy limitada de su función y evolución. El principal obstáculo en la interpretación de las secuencias de nucleótidos es que su mayor parte parece no ser importante. La genómica comparativa es el siguiente paso tras determinar la secuencia de un genoma o una región génica específica y constituye uno de los principales métodos para la anotación funcional de los genomas. Las secuencias funcionales son blancos sobre los cuales actúa la evolución, por ello mediante el alineamiento de secuencias es posible hallar huellas dejadas por presiones selectivas durante millones de años (Miller *et al.*, 2004). A través de la comparación y el análisis de las secuencias de genomas es posible comenzar a descifrar las claves de nuestra historia evolutiva (Alberts 2006).

La gran mayoría de las sustituciones que se producen en las secuencias de ADN en una población o entre especies son selectivamente neutrales (Kimura, 1968; Khaitovich *et al.*, 2006). Las mutaciones en regiones funcionalmente importantes dan lugar a alelos deletéreos, haciendo que aquellos individuos que los hereden mueran sin dejar progenie (Hardison 2003; Miller *et al.*, 2004). De esta forma la selección negativa o purificadora hace que las secuencias funcionales cambien más lentamente que las secuencias que no lo son (o que no realizan funciones significativas), asegurando la retención de dichas funciones. Otras secuencias en cambio quedan a merced de la selección positiva o Darwiniana, muy común después de los eventos de duplicación génica. Estos eventos son de suma importancia para la generación de nuevos genes, donde una de las dos copias del gen queda libre para mutar y especializarse con el fin de llevar a cabo una función diferente, representando un cambio ventajoso para el portador (Alberts 2006; Khaitovich *et al.*, 2006). La especialización se produce de forma gradual, permitiendo que series repetidas de duplicación y divergencia durante millones de años, conduzcan a que un gen de origen a una familia de genes dentro de un único genoma (Alberts 2006), tal es el caso de la familia GH. Esta familia multigénica constituye un buen modelo para estudiar estos y otros aspectos de la evolución de los genomas.

En la evolución humana solamente una pequeña fracción de todos los cambios en la secuencia del ADN entre los seres humanos y los chimpancés es relevante en las

diferencias funcionales entre estas especies, haciendo la identificación de cambios no-neutrales una tarea desalentadora (Khaitovich *et al.*, 2006).

6.1. Obtención de los genes GH/CSH y sus regiones intergénicas.

En el presente trabajo, se obtuvo para el chimpancé la secuencia completa de cinco genes relacionados a GH y se encontró que todos ellos están formados por cinco exones separados por cuatro intrones, en posiciones y tamaños similares a los de los genes del locus humano. Por más que se intentó, no fue posible amplificar la CSH-A reportada previamente (Revol *et al.*, 2004). Coincidentemente, en la información contribuida por el CSGCh tampoco se encontró secuencia alguna que pudiésemos relacionar con dicho gen. Adicionalmente se amplificó el *locus* GH/CSH de chimpancé utilizando múltiples combinaciones de iniciadores que permitieron generar fragmentos con traslapes, los cuales se caracterizaron con enzimas de restricción, clonaron y secuenciaron parcialmente.

Los seis genes previamente reportados para el chimpancé se obtuvieron a partir de ADN genómico (Revol *et al.*, 2004), lo cual puede que haya hecho más complicado el proceso de identificación del número de genes en el *locus*. Debido a la elevada similitud que muestran entre ellos los genes CSH de chimpancé, discriminar entre un alelo y un miembro no-alélico del *locus* puede convertirse en una tarea sumamente difícil. Los resultados obtenidos previamente contrastan en cuanto al número de genes con los obtenidos en este trabajo. Con referencia al número de regiones intergénicas detectadas, previamente fueron reportadas cuatro intergénicas (Revol *et al.*, 2004), lo que apoya los resultados obtenidos en el presente estudio.

Las secuencias obtenidas y las generadas por el CSGCh hicieron posible el ensamblaje final y anotación del *locus* (figura 30). Partiendo de este ensamblaje, se realizaron los análisis de genómica comparativa entre humano y chimpancé descritos a continuación.

6.2. Organización de los genes GH/CSH en el chimpancé.

Para los ensamblajes iniciales realizados en este trabajo se consideró que el *locus* GH del chimpancé estaba formado por seis genes, teniendo en cuenta previas investigaciones de

nuestro laboratorio, cinco estaban secuenciados y el restante sólo contaba con una secuencia parcial (Revol *et al.*, 2004). A pesar de que se logró cubrir en el ensamble la mayor parte del *locus* (figura 11), aun quedaban huecos en regiones intergénicas y dudas en el orden de los genes CSH-A y CSH-C. No obstante, se trabajó sobre el modelo de seis genes tratando de explicar los eventos evolutivos que dieron lugar al *locus* GH humano (Anexo I) y después de muchos análisis se propuso la hipótesis mostrada en el Anexo II para intentar explicar dichos eventos. En esta hipótesis se plantea la delección de un fragmento de aproximadamente 13 kpb, presente en el *locus* del chimpancé, pero ausente en el del humano. Esta delección evolutiva incluye el gen CSH-A y a casi la totalidad de la intergénica larga 2 del chimpancé, de la cual no se tenía secuencias para cubrirla. Múltiples análisis de secuencias permitieron proponer que la delección identificada probablemente hubiese ocurrido por un mecanismo de recombinación entre dos secuencias L1PA3 (LINEs/L1) (Han *et al.*, 2005; Sen *et al.*, 2006).

El hecho de que los análisis no encuadraban bien con el modelo de seis genes para la organización del *locus* GH del chimpancé (Revol *et al.*, 2004), unido a que por más que se intentó no fue posible amplificar el supuesto gen CSH-A, además de cotejar mejor con los datos de secuencias del consorcio, se consideró la posibilidad de que el *locus* pudiese estar formado por cinco genes en vez de seis. Los resultados experimentales obtenidos siempre fueron coincidentes con las predicciones realizadas para el arreglo de cinco genes. Cabe señalar que después de los hallazgos obtenidos en el presente trabajo se descarta la hipótesis inicialmente propuesta para explicar la evolución del *locus* GH humano a partir del *locus* en el chimpancé, pero no descartamos la posibilidad de que un mecanismo de delección similar al que propusimos inicialmente pudo haber originado al *locus* GH del chimpancé, pero esto aún está por investigarse. La secuenciación de esta región génica en otros miembros de los grandes monos nos podría arrojar más luz sobre este particular.

Nuestros hallazgos nos llevaron a concluir que el *locus* GH del chimpancé (figura 30), con una organización casi idéntica a la del humano (Chen *et al.*, 1989) está formado por cinco genes, dos de ellos codifican GHs y los tres restantes CSHs. El primer miembro del *locus* corresponde a la GH-N, seguido por CSH-B, CSH-C, GH-V y CSH-D.

6.3. Comparaciones de los genes GH/CSH de humano y chimpancé.

Las secuencias nucleotídicas obtenidas para los genes del chimpancé se compararon con sus contrapartes en el humano y se estableció la relación filogenética entre ellos. El análisis comparativo de los genes GH-V/CSH de humano y chimpancé respecto a la GH-N (figura 33) revela en ambas especies una tendencia similar de variación en las secuencias nucleotídicas de sus miembros respecto a la GH hipofisiaria. Aunque se pueden observar sutiles diferencias, mientras que en el chimpancé todos los genes CSH exhiben una similitud aproximada de 92% respecto a la GH-N, (lo cual sugiere que estos genes están sometidos a un mismo ritmo de variación), en el humano CSH-1 y CSH-2 muestran este mismo comportamiento con una similitud respecto a la GH-N de alrededor del 92.6%. Finalmente, el pseudogen exhibe un valor ligeramente menor, en congruencia con el hecho de que al no desempeñar un papel funcional significativo, este gen ha quedado libre de cambiar a un ritmo limitado solo por la frecuencia de mutaciones al azar (Alberts 2006).

El análisis comparativo entre los genes GH/CSH ortólogos de humano y chimpancé (figura 34) mostró que todos los genes del chimpancé exhiben una similitud cercana o superior al 98% respecto a su ortólogo en el humano a excepción del gen CSH-B, cuyo valor fue 96%, lo cual era esperado de una comparación con un pseudogen cuya presencia parece ser irrelevante (Gerstein y Zheng, 2006).

Con muy poca divergencia entre las tres copias génicas de CSHs, el gen CSH-C muestra la mayor similitud (98.6%) respecto a su contraparte en el humano. Interesantemente los genes CSH-B y CSH-D del chimpancé mostraron mayor similitud con el gen CSH-1 (98.3% y 98.2% respectivamente) que con su probable ortólogo. Esto podría sugerir que el rumbo evolutivo que están tomando estos genes sea quizás el mismo que hizo que en el humano, las dos copias funcionales de CSHs codifiquen una proteína idéntica (Barrera-Saldaña *et al.*, 1983).

El análisis filogenético se realizó mediante el método de “Neighbor-Joining” (NJ), utilizando la GH del cerdo como grupo fuera (figura 37). El árbol obtenido sugiere que GH-N y GH-V son los únicos genes claramente ortólogos entre el humano y el chimpancé

(González *et al.*, 2006). Por otra parte, para los genes CSH se puede observar que se originaron de un ancestro común, pero en este punto es difícil de establecer entre el humano y el chimpancé la ortología y/o paralogía de estos genes (González *et al.*, 2006). Resultados similares fueron obtenidos usando diferentes métodos filogenéticos para el análisis, tales como Máxima Parsimonia y Likelihood.

6.4. Análisis de las proteínas.

A partir de las secuencias nucleotídicas se determinaron las aminoacídicas y éstas se compararon con sus contrapartes en el humano. En las comparaciones a nivel proteico respecto a la GH-N (figura 35) se observan tendencias similares a las observadas para las secuencias nucleotídicas, excepto nuevamente para el caso del pseudogen, el cual se aleja de la tendencia general que mostró el resto de las CSHs, lo que es justificado por su condición de gen inactivo (Gerstein y Zheng, 2006).

En el humano hay un pseudogen y los dos genes CSH codifican la misma proteína (Barrera-Saldaña *et al.*, 1983), sin embargo, en el chimpancé todos los genes se predicen funcionales y las proteínas codificadas por los genes CSH comparten una similitud superior al 97.9%. Por otro lado se observa una marcada disminución en los valores de similitud para el pseudogen porque este ha acumulado cambios que lo alejan de los demás (Gerstein y Zheng, 2006). El análisis comparativo entre las proteínas ortólogas de humano y chimpancé (figura 36), reveló que entre ellos comparten la misma proteína para la GH-V y que las GH-N de ambas especies exhiben muy alta similitud en sus secuencias aminoacídicas (99.5%), diferenciadas sólo por una sustitución aminoacídica. Estos resultados sugieren que hay una presión selectiva para conservar las GHs (Miller *et al.*, 2004).

En primates la rápida evolución de la GH se produjo de forma simultánea a la de su receptor (GHR), lo cual es sustentado por evidencias que establecen una correspondencia entre los residuos aminoacídicos conservados o sustituidos en la hormona con respecto a su receptor. Un claro ejemplo es que en MNM con respecto a no-primates y prosimios, se ha observado una sustitución del residuo His por el residuo Asp en la posición 171 de la GH,

la cual parece haber dictado en MVM un cambio de Leu por Arg en la posición 43 del GHR. Estos cambios que tuvieron lugar en aminoácidos que intervienen en la unión hormona-receptor han sido conservados por la evolución en órdenes superiores (Souza *et al.*, 1995; Liu *et al.*, 2001) y se traducen en la afinación de esta interacción molecular. A nivel fisiológico la evolución de esta interacción se evidencia por la capacidad de la hGH de unirse al GHR de especies filogenéticamente inferiores y no a la inversa (Liu *et al.*, 2001).

El significado potencial de los cambios aminoacídicos de todas las proteínas respecto a la GH-N del humano, puede ser más revelador si se circunscribe a las posiciones o residuos aminoacídicos involucrados en la unión al receptor. Como se puede deducir de las comparaciones entre estas (figura 32), mientras que las GH-N y GH-V del chimpancé muestran uno o ningún cambio respecto a sus contrapartes en el humano (cuadro 8), los genes CSH de este primate muestran un mayor número de cambios, incluido uno en la posición 164 de las referidas regiones de importancia para la afinidad por el receptor (Liu *et al.*, 2001). Este cambio presente en los tres genes CSH de chimpancé es una sustitución de un residuo tirosina por histidina que también está presente en todos los genes CSH del mono Rhesus (datos no mostrados) (González *et al.*, 2006). Es posible que este residuo aminoacídico no impida la unión al receptor, aunque quizás sí afecte la afinidad con la que se da (por su contribución a la conformación espacial del sitio de unión al receptor) y en consecuencia las potencias relativas de las actividades somatogénicas y lactogénicas (Duda y Brooks, 2003). Se requieren más investigación sobre la estructura y función de estas hormonas en primates, para intentar entender la naturaleza de los cambios observados en estas especies.

6.5. Análisis de los elementos reguladores.

Entender la regulación de los genes ha sido y sigue siendo uno de los desafíos principales para la comunidad de biólogos moleculares. La regulación génica es mediada por una variedad de cortas secuencias de ADN llamadas elementos reguladores, que incluyen sitios de unión a factores de la transcripción. Un primer paso en el avance hacia la comprensión de la regulación de los miembros del locus GH es la identificación de los elementos

reguladores presentes en éste. El segundo paso es la identificación de huellas en el ADN que revelen acontecimientos evolutivos a los que este fue sometido. La llamada huella digital filogenética se basa en la observación de los elementos reguladores que están bajo presión selectiva, que hace que ellos evolucionen a una tasa más lenta que la secuencia no funcional circundante. Así, comparando regiones reguladoras ortólogas de especies relacionadas, la huella digital filogenética predice las subsecuencias altamente conservadas que podrían funcionar como elementos reguladores potenciales (Prakash y Tompa, 2005).

6.6. Análisis del promotor proximal.

Antes del codón de inicio de los genes GH/CSH del humano han sido identificadas dentro del promotor varias secuencias reguladoras y el análisis de estas, en el chimpancé, ha revelado que todas están altamente conservadas entre ambas especies (figura 39). La caja TATA está conservada en todos los casos. Los elementos Pit-1 proximal y distal se encuentran altamente conservados en todos los genes, incluido el pseudogen humano, aunque se ha demostrado que Pit-1 no participa en la expresión de los genes GH/CSH placentarios del humano (Jiang *et al.*, 1995).

Resulta interesante que la secuencia del elemento Pit-1 proximal para la GH-V de humano y chimpancé exhibe variaciones sustanciales respecto a la del resto de los genes, pero llama la atención que estas variaciones parecen retenerse en ambas especies, incluso en el promotor de la GH-V del orangután (datos no mostrados), sugiriendo un importante papel en la expresión de dichos genes (Miller *et al.*, 2004; Prakash y Tompa, 2005). Es muy probable que estos cambios que se retienen en el promotor de la GH-V constituyan un sitio de unión a factores de transcripción que están involucrados en la regulación de la expresión de dicho gen (Prakash y Tompa, 2005). Para la comprensión del significado biológico de estos cambios se requieren ensayos de expresión en líneas celulares de placenta utilizándose un gen reportero para su evaluación.

Los elementos de respuesta a hormona tiroidea (TRE) han sido identificados en la región promotora del gen de la GH y de los genes CSH placentarios humanos (Leidig *et al.*, 1992). Estos TREs se unen al receptor de la hormona tiroidea, pero solo inducen un aumento de la

transcripción de los genes CSH placentarios. Interesantemente, en esta región la secuencia de los genes CSH-B y CSH-D de chimpancé son más similares a CSH-1 y CSH-2 de humano que al resto de los genes. Las posiciones -49 y -47 identificadas como cruciales para la función de estos elementos (Leidig *et al.*, 1992) se hayan ocupadas por los nucleótidos **G** y **A** respectivamente, como ocurre en los genes CSH humanos y no en la GH. Para el caso del gen CSH-C de chimpancé, en esta región su secuencia es similar a la de la GH humana por lo cual se predice una falta de funcionalidad para dicho elemento. Por otro lado, este elemento para el gen GH-V de humano y chimpancé no es funcional, sin embargo los cambios que presenta se encuentran retenidos en ambas especies y también en el orangután (datos no mostrados), sugiriendo que esta secuencia puede jugar un papel funcional en la regulación de la expresión de dicho gen en placenta (Prakash y Tompa, 2005).

El sitio de unión del iniciador InrE, ubicado entre las posiciones -15 y -1, está conservado en todas las secuencias, lo que reafirma su importancia en la regulación de estos genes. InrE induce la transcripción de los genes CSH en la placenta, pero no para los genes GH. La secuencia del sitio InrE se requiere para la actividad eficiente del promotor y máxima del potenciador (Jiang *et al.*, 1995).

6.7. El elemento P.

El elemento P se encontró en cada una de las regiones intergénicas del chimpancé. En el genoma humano el elemento P está situado en las posiciones -2,439; -2,386; -2,287 y -2,690 de los genes CSH-L, CSH-A, GH-V y CSH-B, respectivamente. En el chimpancé está en las posiciones -2,258; -2,249; -2,222 y -2,515 de los genes CSH-B, CSH-C, GH-V y CSH-D, respectivamente. Por otra parte, el alto porcentaje de similitud entre los *loci* GH del humano y del chimpancé (figura 40), sugiere que comparten el mismo mecanismo de control de la transcripción de los genes CSHs (Norquay *et al.*, 2006).

6.8. El potenciador.

El potenciador únicamente se encontró río abajo de los genes CSH del chimpancé. Interesantemente, los potenciadores de los genes CSH-B y CSH-C del chimpancé poseen la

inserción “AGAA” río arriba del dominio DF-2 (figura 41). Dicha inserción se ha descrito para el humano en los potenciadores de los genes CSH-L y CSH-A como de interferencia con su función (Jacquemin *et al.*, 1994). Estos datos sugieren mecanismos similares de la regulación diferencial de los genes CSH, entre el humano y el chimpancé.

En el humano se sabe que la actividad del potenciador está mediada principalmente por DF3 y DF4, sin embargo los potenciadores de CSH-L y CSH-A muestran débil actividad debido a dos mutaciones puntuales en sus respectivos sitios DF3 (Jacquemin *et al.*, 1996). En la figura 41 se observa que dichos sitios en los potenciadores de los genes CSH-B y CSH-C de chimpancé tienen idéntica secuencia, por lo cual son también portadores de estas dos mutaciones que le predicen débil actividad. El potenciador del gen CSH-D solo presenta una de estas mutaciones, sugiriendo que éste probablemente no tenga más actividad que el potenciador del gen CSH-B del humano, pero si más que el resto de los potenciadores de los genes CSH de ambas especies.

6.9. Elemento de respuesta a glucocorticoides

Se ha reportado la presencia de un elemento de respuesta a glucocorticoides (GRE) en el primer intrón del gen de la GH-N humana (Slater *et al.*, 1985). En la figura 42 se muestra un alineamiento de las secuencias nucleotídicas de los genes GH/CSH de humano y chimpancé, el GRE fue identificado en el alineamiento tomando como referencia el reporte previo de su secuencia para la hGH. Este elemento se encuentra conservado en la GH-N y GH-V de chimpancé, nótese que también lo está en la GH-V humana.

Se ha descrito ampliamente la capacidad reguladora transcripcional que tienen los glucocorticoides a través de los elementos de respuesta a los mismos (GRE). La identificación de un segmento de ADN que tiene la propiedad de un GRE no sólo se encontró en el primer intrón del gen de la GH-N humana, sino además en idéntica posición en la GH-V humana, GH-N y GH-V de chimpancé, pero no en los genes GH del marmoset (Wallis y Wallis, 2002).

6.10. Regiones intergénicas

De igual forma que como ocurre con los genes, las regiones intergénicas se han estado conservando. Como se muestra en la figura 38 todas las regiones intergénicas exhiben una similitud superior al 96.9%, lo que sugiere que en ellas se encuentran elementos reguladores de suma importancia en la expresión de los genes miembros del *locus*, incluyendo aquellos responsables de la especificidad tisular de expresión (Consortio del Proyecto ENCODE, 2004; Prakash y Tompa, 2005).

En las regiones intergénicas largas, pero no en las cortas, se identifican regiones que presentan similitud con el gen TCAM que flanquea al *locus* GH en su extremo 3'. Se presume que estas secuencias se originaron durante los eventos evolutivos a partir de dicho gen y posteriormente fueron retenidas durante los eventos de duplicación que tuvieron lugar en el *locus* GH (González *et al.*, 2006).

6.11. Elementos repetitivos.

Los elementos *Alu* han contribuido claramente a la evolución del genoma humano y en los primates han mostrado una importante actividad en los eventos evolutivos (Kehrer-Sawatzki y Cooper, 2007). En el *locus* GH puede ser que los elementos *Alu* no solo hayan jugado un importante papel en la generación del *locus* (Chen *et al.*, 1989), sino también, al menos en el caso del elemento *Alu* Sx situado justamente detrás de los genes hGH-N y hGH-V, en la regulación transcripcional de estos genes. En el humano ha sido demostrado (Trujillo *et al.*, 2006) que este elemento *Alu* Sx contiene un silenciador compuesto por cuatro dominios de interacción que incluyen un dominio represor (CRD), un dominio antisilenciador (ASE) que contiene elementos que median la actividad silenciadora independiente de la orientación y dos dominios flanqueantes que modulan la actividad silenciadora de manera dependiente del CRD. El silenciador es dependiente del promotor GH y su función es específica para las células que expresen hGH-N o hGH-V (Trujillo *et al.*, 2006).

Las regiones intergénicas cortas y largas del chimpancé presentan un patrón similar al del humano en cuanto a la distribución de los elementos *Alu*, pero resulta aun más interesante

que el silenciador contenido en la secuencia *Alu* Sx que ha sido reportado para el humano, también esté presente detrás de los genes GH-N y GH-V del chimpancé, sugiriendo una conservación de este mecanismo de regulación transcripcional entre ambas especies. Aunque es muy posible que esa secuencia reguladora sea una adquisición evolutiva relativamente reciente (Trujillo *et al.*, 2006).

6.12. El valor de la genómica comparativa.

Recientes análisis de genómica comparativa indican que las marcadas diferencias entre los humanos modernos y los chimpancés se deben más a cambios en la regulación génica que a las diferencias en el contenido génico. Casi semanalmente se publican noticias de genes y elementos reguladores identificados que parecen experimentar una fuerte selección positiva a lo largo del linaje humano, los *loci* involucrados han sido de particular interés (Bradley, 2008). Investigaciones tales como el proyecto del genoma del hombre de Neanderthal prometen que los análisis moleculares continuaran proporcionando nuevas visiones acerca de nuestra evolución (Green *et al.*, 2006).

La secuencia del genoma de un individuo necesita solamente ser determinada una vez, sin embargo los niveles de expresión cambian en las diferentes etapas de desarrollo, en los diversos tejidos, y diversos ambientes (Khaitovich *et al.*, 2006). El presente trabajo señala cambios sutiles en los genes o incluso la ausencia de estos entre ambos *loci*, aunque resultará importante evaluarlos, también será menester complementarla con estudios de expresión de las mismas, ya sea en cultivo celular o de ser posible, y mejor aún, en tejidos como la placenta.

6.12. Perspectivas.

Las clonas que contenían los plásmidos portadores de los genes GH/CSH de chimpancé e incorporados los sitios de restricción (*KpnI* en el 5' y *EcoRV* en el 3') para permitir su posterior subclonación unidireccional en el vector de expresión pcDNA 3.1 hygro (+) fueron proporcionadas a la Unidad de Biotecnología Médica para su subsecuente determinación del potencial codificante y funcional. De estas investigaciones ya se descubrió que el gen GH-N no ha perdido funciones conforme se sube en la escala

evolutiva y el locus adquiere genes del tipo CSH y GH-V. Estudios similares seguro arrojaría luz sobre el papel de los otros miembros de la familia, así como también del origen de su especificidad tisular y temporal.

CAPÍTULO VII: CONCLUSIONES

- 1.- Se determinó la organización y composición del *locus* GH de chimpancé, validándose ésta por diferentes enfoques de laboratorio.
- 2.- El *locus* GH del chimpancé comprende dos genes GHs y tres CSHs, separados por cuatro regiones intergénicas, que de acuerdo a su tamaño se clasifican en dos cortas (6 y 6.5 kpb) y dos largas (13 kpb).
- 3.- Comparaciones de las secuencias descubiertas en el *locus* GH del chimpancé, en relación a su contraparte humana, mostraron valores muy elevados para cada uno de los elementos analizados.
- 4.- Se comprobó la ortología de los genes GH de ambos *loci* y se infiere un mayor grado de conservación en estos con respecto a los CSH.
- 5.- El promotor y otros elementos clave de la regulación génica, están conservados en los *loci* GH del humano y chimpancé.

CAPÍTULO VIII: BIBLIOGRAFIA

- Fleischmann RD, MD Adams, O White, RA Clayton, EF Kirkness, AR Kerlavage, CJ Bult, JF Tomb, BA Dougherty, JM Merrick. *Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.* Science, Jul 1995; 269: 496 – 512
- Genome Online Database GOLD. News 2008. Consulta 22 de Mayo de 2008. http://www.genomesonline.org/gold_news.htm
- Roberts, Leslie, R. John Davenport, Elizabeth Pennisi, and Eliot Marshall. Science. *A History of the Human Genome Project.* 2001: Vol. 291. no. 5507, p. 1195.
- S. Choudhuri. *The Path from Nuclein to Human Genome: A Brief History of DNA with a Note on Human Genome Sequencing and Its Impact on Future Research in Biology.* Bulletin of Science Technology Society, 2003. 23, 360-367.
- Facts About Genome Sequencing.* Octubre 2006. Consulta 22 de Mayo de 2008. http://www.ornl.gov/sci/techresources/Human_Genome/faq/seqfacts.shtml#complete
- Spencer, Geoff. *International Consortium Completes Human Genome Project*, NHGRI. Consulta 23 de Mayo de 2008. <http://www.genome.gov/11510908>
- Proyecto del Genoma Humano. Instituto Nacional de Medicina Genómica.* Mexico. http://inmegen.gob.mx/index.php?option=com_content&task=view&id=88&Itemid=213&mo=1&yr=2007&lang=es Consulta 25 de Mayo de 2008.
- Large Scale Genome Sequencing Projects. 1000 Genomes, A Deep Catalog of Human Genetic Variation.* GOLD, 2008. Consulta 26 de Mayo de 2008. <http://www.1000genomes.org/files/1000Genomes-NewsRelease.pdf>
- Rouze P, Pavy N, Rombauts S. (1999) Genome annotation: which tools do we have for it? *Curr Opin Plant Biol.* 2:90-95.
- Stein L. (2001) Genome annotation: from sequence to biology. *Nat Rev Genet.* 2:493-503.
- Davidson, Eric H., Jonathan P. Rast, Paola Oliveri, Andrew Ransick, Cristina Calestani, Chiou-Hwa Yuh, Takuya Minokawa, Gabriele Amore, Veronica Hinman, César Arenas-Mena, Ochan Otim, C. Titus Brown, Carolina B. Livi, Pei Yun Lee, Roger Revilla, Alistair G. Rust, Zheng jun Pan, Maria J. Schilstra, Peter J. C. Clarke, Maria I. Arnone, Lee Rowen, R. Andrew Cameron, David R. McClay, Leroy Hood, and Hamid Bolouri. *A Genomic Regulatory Network for Development.* Science, Mar 2002; 295: 1669.
- Koonin EV. (2001) Computational genomics. *Curr Biol.* 11:R155-158.
- Salzberg SL. (2003) Genomics: Yeast rises again. *Nature.* 423:233-234.

- Zuckermandl, Emile, Linus Pauling. *The Molecular Evolutionary Clock*. Journal of the History of Biology. Springer Netherlands. Vol 31. 1965
- Análisis de genomas. Métodos para la predicción y anotación de la función de las proteínas*. Tesis doctoral, Universidad Autónoma de Madrid, Abascal, F., 2003
- Lesk, Arthur M., Michael Levitt, and Cyrus Chothia. *Alignment of the amino acid sequences of distantly related proteins using variable gap penalty*. Protein Eng, Oct 1986; 1: 77 - 78.
- Henikoff, Steven, Elizabeth A. Greene, Shmuel Pietrokovski, Peer Bork, Teresa K. Attwood, and Leroy Hood *Gene Families: The Taxonomy of Protein Paralogues and Chimeras*. Science, Oct 1997; 278: 609.
- Horgan, Douglas J., H. Ohno, Thomas P. Singer, and John E. Casida. *Studies on the Respiratory Chain-linked Reduced Nicotinamide Adenine Dinucleotide Dehydrogenase. XV. INTERACTIONS OF PIERICIDIN WITH THE MITOCHONDRIAL RESPIRATORY CHAIN*. J. Biol. Chem., Nov 1968; 243: 5967 - 5976.
- Watts RL, Watts DC. (1968) Gene duplication and the evolution of enzymes. *Nature*. 217:1125-1130.
- Ohta T. (1989) Role of gene duplication in evolution. *Genome*. 31:304-310.
- Lang D, Thoma R, Henn-Sax M, Sterner R, Wilmanns M. (2000) Structural evidence for evolution of the beta/alpha-barrel scaffold by gene duplication and fusion. *Science*. 289:1546-1550.
- 21.- Nowell, Susan A., Julian E. A. Leakey, Jason F. Warren, Nicholas P. Lang, and Lynn T. Frame. *Identification of Enzymes Responsible for the Metabolism of Heme in Human Platelets*. J. Biol. Chem., Dec 1998; 273: 33342.
- Glazko, Galina V., Masatoshi Nei *Estimation of Divergence Times for Major Lineages of Primate Species*. Mol. Biol. Evol., Mar 2003; 20: 424 - 434.
- Mikkelsen, Tarjei S. et al., *Initial sequence of the chimpanzee genome and comparison with the human genome*. Nature, 2005. 437(7055): p. 69-87.
- Wilson, A. C., G. L. Bush, S. M. Case, and M.-C. King. *Social Structuring of Mammalian Populations and Rate of Chromosomal Evolution*. PNAS, Dec 1975; 72: 5061 - 5065.
- MV, Olson; Varki A. *The chimpanzee genome - A bittersweet celebration*. Journal : Editorial Material. SCIENCE 305 (5681): 191-192
- Enard, Wolfgang; Svante Pääbo, Khaitovich, Philipp, Bjoern Muetzel, Xinwei She, Michael Lachmann, Ines Hellmann, Janko Dietzsch, Stephan Steigle, Hong-Hai Do, Gunter Weiss, Florian Heissig, Thomas Arendt, Kay Nieselt-Struwe, Evan E. Eichle.

- Regional Patterns of Gene Expression in Human and Chimpanzee Brains.* Genome Res., Aug 2004; 14: 1462 - 1473.
- Varki, Ajit; Tasha K. Altheide. *Comparing the human and chimpanzee genomes: Searching for needles in a haystack.* Genome Res., Dec 2005; 15: 1746 - 1758.
- Khaitovich, P., et al., *Evolution of primate gene expression.* Nat Rev Genet, 2006. 7(9): p. 693-702.
- New findings from the ENCODE project challenge established views on human genome.* 2007. Consulta Junio 1 de 2008. http://www.cbse.ucsc.edu/news/2007/06/13/encodeposter_nature0607sm.jpg
- Evolución de los primates.* Modificado de Great Ape Trust. Consulta Mayo 28 de 2008. <http://www.greatapetrust.org/primates/index.php>.
- Wallis, O.C., Z. Maniou, and M. Wallis, *Cloning and characterization of the gene encoding growth hormone in finback whale (Balaenoptera physalus).* Gen Comp Endocrinol, 2005. 143(1): p. 92-7.
- Chen, E.Y., et al., *The human growth hormone locus: Nucleotide sequence, biology, and evolution.* Genomics, 1989. 4(4): p. 479-497.
- Revol de Mendoza, A., et al., *Expansion and divergence of the GH locus between spider monkey and chimpanzee.* Gene, 2004. 336(2): p. 185-193.
- Golos ,TG; M Durning, JM Fisher, and PD Fowler. *Cloning of four growth hormone/chorionic somatomammotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta.* Endocrinology, Oct 1993; 133: 1744.
- Ye, Chun, e.a., *Molecular evolution of growth hormone gene family in old world monkeys and hominoids.* Gene, 2005. 350: p. 183-192.
- Wallis, O.C., et al., *Polymorphism of the growth hormone gene of red deer (Cervus elaphus).* Gen Comp Endocrinol, 2006. 146(2): p. 180-5.
- Li, Y., et al., *Independent origin of the growth hormone gene family in New World monkeys and Old World monkeys/hominoids.* 2005. p. 399-409.
- George DL, JA Phillips III, U. Francke, PH Seeburg 1981. The genes from growth hormone and chorionic somatomammotropin are on the long arm of the chromosome 17 in region q21-q22. Human. Genet. 57:138-141
- Barrera-Saldana, H.A., Seeburg, P.H., Saunders, G.F., 1983. Two structurally different genes produce the same secreted human placental lactogen hormone. J Biol Chem 258, 3787-3793.

- Reséndez-Pérez D, Ramírez-Solis R. Varela-Echavarria A., Martínez- Rodríguez HG, Barrera-Saldaña HA. 1990. Coding potential of transfected human placental genes. *Nucleic Acids Res*, 18: 4665-4670
- Barrera-Saldaña, HA; Harper ME, GF Saunders. 1982. Chromosome localization of the human placental lactogene-growth hormone gene cluster to 17q22-24. *Am. J. Human . Genet.* 34,227-234
- Handwerger, S., Freemerk, M., 2000. The roles of placental growth hormone and placental lactogen in the regulation of human fetal growth and development. *J Pediatr Endocrinol Metab* 13, 343-356.
- Frankenne F, F. Rentier-Delrue, ML, Scippo, J. Martical, G. Hennen 1987. Expression of the growth hormone variant gene in human placenta. *J. Clin. Endocrin. Metab.* 64:635-637
- Cooke N.E. Ray J. Emery, JG, Liebhaber A. 1988. Two distinct species of human growth hormone-variant mRNA in the human placenta the expression of novel growth hormone proteins . *J. Biol. Chem* 263:9001-9006F
- Verhaeghe, J., Pintiaux, A., Van Herck, E., Hennen, G., Foidart, J.M., Igout, A., 2002. Placental GH, IGF-I, IGF-binding protein-1, and leptin during a glucose challenge test in pregnant women: relation with maternal body weight, glucose tolerance, and birth weight. *J Clin Endocrinol Metab* 87, 2875-2882.
- Chellakooty, M., Vangsgaard, K., Larsen, T., Scheike, T., Falck-Larsen, J., Legarth, J., Andersson, A.M., Main, K.M., Skakkebaek, N.E., Juul, A., 2004. A longitudinal study of intrauterine growth and the placental growth hormone (GH)-insulin-like growth factor I axis in maternal circulation: association between placental GH and fetal growth. *J Clin Endocrinol Metab* 89, 384-391.
- Frankenne F, F. Rentier-Delrue, ML, Scippo, J. Martical, G. Hennen 1987. Expression of the growth hormone variant gene in human placenta. *J. Clin. Endocrin. Metab.* 64:635-637
- Karabulut, A.K., Layfield, R., Pratten, M.K., 2001. Growth promoting effects of human placental lactogen during early organogenesis: a link to insulin-like growth factors. *J Anat* 198, 651-662.
- Ruiz Parra, A.I. Metabolismo en el embarazo entre lo elástico y lo actual. <http://encolombia.com/obstetricia50399editorial.html>.
- Walker W.H., Fitzpatrick S.L., Barrera-Saldaña H.A., Reséndez-Perez D., Saunders G.F., 1991. The human placental lactogen genes: Structure, function, evolution and transcriptional regulation. *Endocrine Reviews* 12, 316-328.

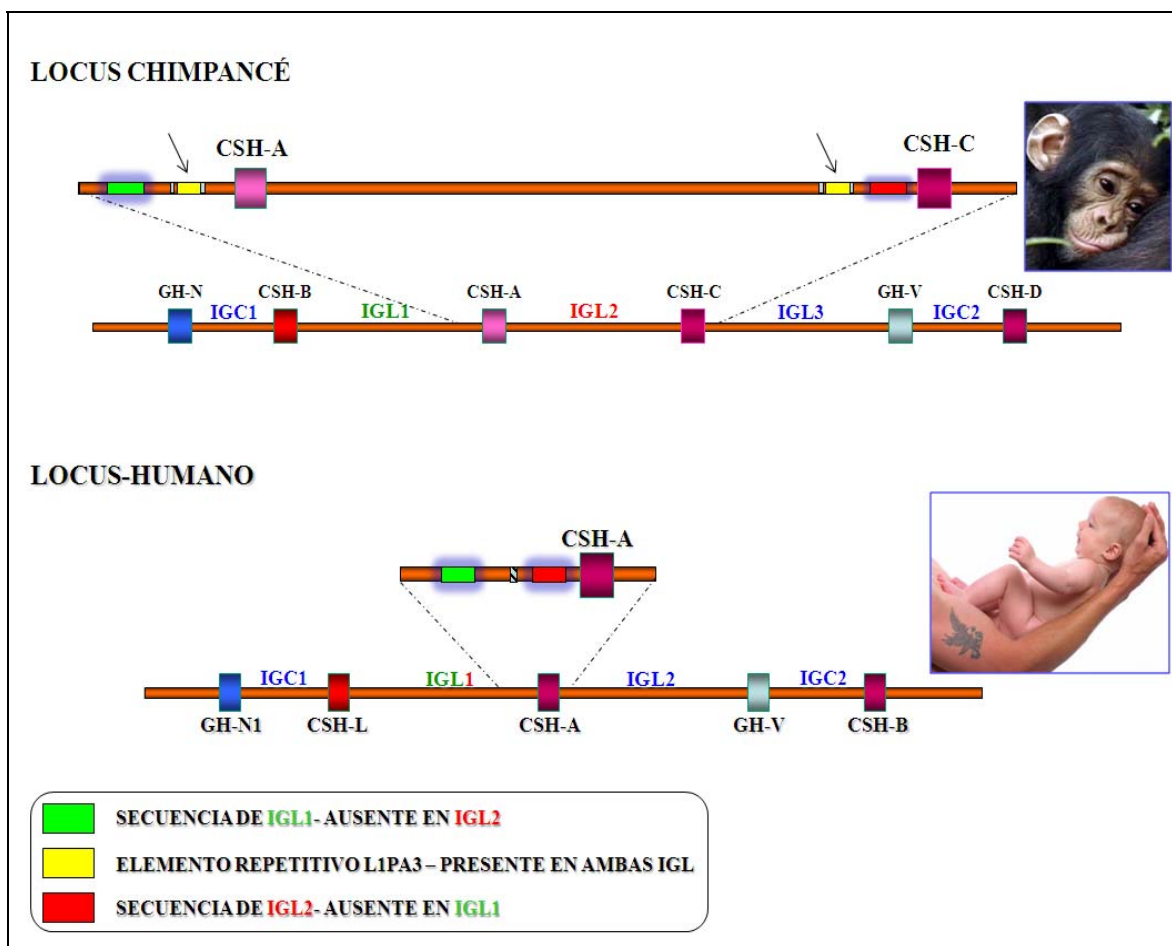
- Jacquemin, P., et al., *Characterization of a single strong tissue-specific enhancer downstream from the three human genes encoding placental lactogen*. Mol Cell Biol, 1994. 14(1): p. 93-103.
- Kimura, A.P., S.A. Liebhaber, and N.E. Cooke, *Epigenetic modifications at the human growth hormone locus predict distinct roles for histone acetylation and methylation in placental gene activation*. Mol Endocrinol, 2004. 18(4): p. 1018-32.
- Shewchuk, B.M., et al., *A Single Base Difference between Pit-1 Binding Sites at the hGH Promoter and Locus Control Region Specifies Distinct Pit-1 Conformations and Functions*. 2006. p. 6535-6546.
- Karin M., Theill, L., Castrillo, J.S., McCormick, A., Brady, H. 1990. Cell type specific expression of the growth hormone gene and its control by GHF-1. Folia endocrinol. 66. 1205-1220
- Ingraham, H.A., Chen, R.P., Mangalam, H.J., Elsholtz, H.P., Flynn, S.E., Lin, C.R., Simmons, D.M., Swanson, L., Rosenfeld, M.G., 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell 55, 519-529.
- Norquay, L.D., et al., *A member of the nuclear factor-1 family is involved in the pituitary repression of the human placental growth hormone genes*. Biochem J, 2001. 354(Pt 2): p. 387-95.
- Nachtigal, M.W., Nickel, B.E., Cattini, P.A., 1993. Pituitary-specific repression of placental members of the human growth hormone gene family. A possible mechanism for locus regulation. J Biol Chem 268, 8473-8479.
- Jiang, S.W., Shepard, A.R., Eberhardt, N.L., 1995. An initiator element is required for maximal human chorionic somatomammotropin gene promoter and enhancer function. J Biol Chem 270, 3683-3692.
- Rogers BL, MG, Sobnosky, GF, Saunders. 1986. Transcriptional enhancer within the human placental lactogen and growth hormone multigene cluster. Nucleic Acids Res. 14: 7647-7659
- Jacquemin, P., Oury, C., Peers, B., Morin, A., Belayew, A., Martial, J.A., 1994. Characterization of a single strong tissue-specific enhancer downstream from the three human genes encoding placental lactogen. Mol Cell Biol 14, 93-103.
- Sambrook Joseph., R.W.D., 2001. Molecular Cloning A Laboratory Manual. Cold Spring Harbor, New York.
- Huang, X. and Madan, A. (1999) CAP3: A DNA sequence assembly program. Genome Res., 9, 868-877.

- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4, 406-425.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5, 150-163.
- Burge, C. and Karlin, S. (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78-94
- Burge, C. B. (1998) Modeling dependencies in pre-mRNA splicing signals. In Salzberg, S., Searls, D. and Kasif, S., eds. *Computational Methods in Molecular Biology*, Elsevier Science, Amsterdam, pp. 127-163.
- Burge, C. B. and Karlin, S. (1998) Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.* 8, 346-354.
- Liu, J.C., Makova, K.D., Adkins, R.M., Gibson, S., Li, W.H., 2001. Episodic evolution of growth hormone in primates and emergence of the species specificity of human growth hormone receptor. *Mol Biol Evol* 18, 945-953.
- Green, P.; Gordon, D., Abajian, C., 1998. *Consed: a graphical tool for sequence finishing*. *Genome Res* 8, 195-202.
- Goffin, V., Shiverick, K.T., Kelly, P.A., Martial, J.A., 1996. Sequence-function relationships within the expanding family of prolactin, growth hormone, placental lactogen, and related proteins in mammals. *Endocr Rev* 17, 385-410.
- Peterson, F.C. and Brooks, C.L. Identification of a Lactogenic Motif in Human Growth Hormone. *Journal of Biological Chemistry* 212: 21444-21448 (1997).
- K.M. Duda and C.L. Brooks. *Differential effects of zinc on functionally distinct human growth hormone mutations*
- Tatiana A. Tatusova, Thomas L. Madden (1999), "Blast 2 sequences - a new tool for comparing protein and nucleotide sequences", *FEMS Microbiol Lett.* 174:247-250.
- Burge, C. and Karlin, S. (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78-94.
- Burge, C. B. (1998) Modeling dependencies in pre-mRNA splicing signals. In Salzberg, S., Searls, D. and Kasif, S., eds. Computational Methods in Molecular Biology, Elsevier Science, Amsterdam, pp. 127-163.
- Burge, C. B. and Karlin, S. (1998) Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.* 8, 346-354.

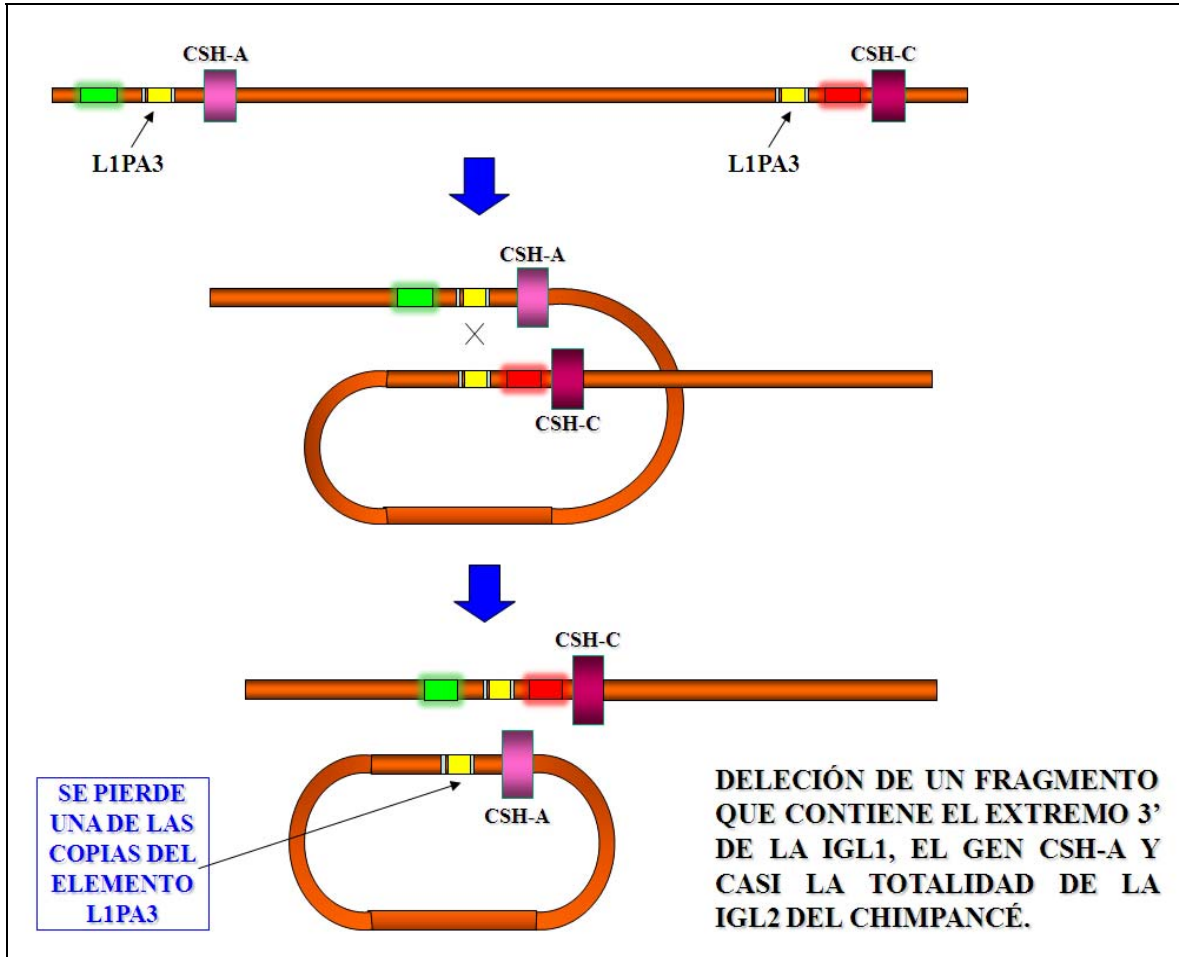
- Huang, X. and Madan, A. (1999) CAP3: A DNA sequence assembly program. *Genome Res.*, 9, 868-877.
- Evolution of primate gene expression. Philipp Khaitovich, Wolfgang Enard, Michael Lachmann & Svante Pääbo. Review. *Nature Reviews Genetics* 7, 693-702 (September 2006)
- Discovery of regulatory elements in vertebrates through comparative genomics. Prakash A y Tompa M. *Nat Biotechnol.* 2005 Oct; 23(10):1249-56.
- Walker WH, Fitzpatrick SL, Barrera-Saldaña HA, Resendez-Perez D, Saunders GF. *Endocr Rev.* 1991 Nov., 12(4):316-28. The human placental lactogen genes: structure, function, evolution and transcriptional regulation.
- A History of the Human Genome Project. Roberts *et al.*, 2001 *Science*: Vol. 291. no. 5507, p. 1195.
- [Genome OnLine Database, (GOLD). News, 2008. http://www.genomesonline.org/gold_news.htm]
- The Path from Nuclein to Human Genome: A Brief History of DNA with a Note on Human Genome Sequencing and Its Impact on Future Research in Biology. S. Choudhuri (2003) *Bulletin of Science Technology Society* 23, 360-367.
- Facts About Genome Sequencing. http://www.ornl.gov/sci/techresources/Human_Genome/faq/seqfacts.shtml#complete
- International Consortium Completes Human Genome Project. Geoff Spencer, NHGRI. <http://www.genome.gov/11510908>
- Proyecto del Genoma Humano. Instituto Nacional de Medicina Genomica. Mexico. Consulta 25 de Mayo de 2008. http://innegen.gob.mx/index.php?option=com_content&task=view&id=88&Itemid=213&mo=1&yr=2007&lang=es
- (GOLD. Large Scale Genome Sequencing Projects. 1000 Genomes, A Deep Catalog of Human Genetic Variation. 2008. <http://www.1000genomes.org/files/1000Genomes-NewsRelease.pdf>)
- (Análisis de genomas. Métodos para la predicción y anotación de la función de las proteínas. Tesis doctoral, Universidad Autónoma de Madrid, Abascal, F., 2003)
- [New findings from the ENCODE project challenge established views on human genome. 2007. (http://www.cbse.ucsc.edu/news/2007/06/13/encodeposter_nature0607sm.jpg)]
- (Ruiz Parra, A.I. Metabolismo en el embarazo entre lo clásico y lo actual. <http://encolombia.com/obstetricia50399editorial.htm>).
- (Consorcio del Proyecto ENCODE, 2004). "The ENCODE (ENCyclopedia Of DNA Elements) Project." *Science* 306(5696): 636-40.

- Birney, E., et al., (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." *Nature* 447(7146): 799-816.
- Lander, E. S., L. M. Linton, et al. (2001). "Initial sequencing and analysis of the human genome." *Nature* 409(6822): 860-921.
- Hardison, R. C. (2003). "Comparative genomics." *PLoS Biol* 1(2): E58.
- Jacquemin, P., E. Alsat, et al. (1996). "The enhancers of the human placental lactogen B, A, and L genes: progressive activation during in vitro trophoblast differentiation and importance of the DF-3 element in determining their respective activities." *DNA Cell Biol* 15(10): 845-54.
- Han, K., S. K. Sen, et al. (2005). "Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages." *Nucleic Acids Res* 33(13): 4040-52.
- Sen, S. K., K. Han, et al. (2006). "Human genomic deletions mediated by recombination between Alu elements." *Am J Hum Genet* 79(1): 41-53.
- Gerstein, M. and D. Zheng (2006). "The real life of pseudogenes." *Sci Am* 295(2): 48-55.
- Kimura, M. (1968). "Evolutionary rate at the molecular level." *Nature* 217(5129): 624-6.
- Norquay, L. D., X. Yang, et al. (2006). "Hepatocyte nuclear factor-3alpha binding at P sequences of the human growth hormone locus is associated with pituitary repressor function." *Mol Endocrinol* 20(3): 598-607.
- Trujillo, M. A., M. Sakagashira, et al. (2006). "The human growth hormone gene contains a silencer embedded within an Alu repeat in the 3'-flanking region." *Mol Endocrinol* 20(10): 2559-75.
- Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." *J Mol Biol* 215(3): 403-10.
- Kehrer-Sawatzki, H. and D. N. Cooper (2007). "Understanding the recent evolution of the human genome: insights from human-chimpanzee genome comparisons." *Hum Mutat* 28(2): 99-130.
- Sakatani, S., R. Takahashi, et al. (2000). "Structure, expression, and conserved physical linkage of mouse testicular cell adhesion molecule-1 (TCAM-1) gene." *Genome* 43(6): 957-62.
- Shewchuk, B. M., Y. Ho, et al. (2006). A Single Base Difference between Pit-1 Binding Sites at the hGH Promoter and Locus Control Region Specifies Distinct Pit-1 Conformations and Functions. 26: 6535-6546.
- Green, R. E., J. Krause, et al. (2006). "Analysis of one million base pairs of Neanderthal DNA." *Nature* 444(7117): 330-6.

CAPÍTULO VIII: ANEXOS



Anexo 1. Diferencia principal entre los *loci* GH de humano y chimpancé. Análisis basado en el ensamblaje inicial obtenido en este trabajo para el *locus* GH del chimpancé considerando que su organización se ajustaba a un modelo de seis genes. Ampliada se representa la región crítica considerada para proponer como hipótesis un evento evolutivo que implicaba una delección en el *locus* GH del chimpancé y cuyo resultado explicaba las diferencias encontradas en dichos análisis con respecto al humano.



Anexo 2. Hipótesis propuesta en base al modelo de seis genes inicialmente considerado para el *locus* GH del chimpancé. Se plantea como hipótesis un evento evolutivo que implicaba la deleción, en el *locus* GH del chimpancé, de un fragmento de aproximadamente 13 kpb que incluía al gen CSH-A y casi la totalidad de la intergénica larga dos de este primate. Se propuso que la deleción probablemente hubiese ocurrido por un mecanismo de recombinación homóloga entre dos copias de un elemento L1PA3.